高等教育质量工程信息技术系列示范教材

# 信息系统安全教程

（第3版）

张基温 编著



清华大学出版社

高等教育质量工程信息技术系列示范教材

# 信息系统安全教程
# （第3版）

张基温　编著

清 华 大 学 出 版 社
北　京

## 内 容 简 介

本书从应用的角度介绍计算机信息系统安全技术。全书按照“威胁—防护—管理”的思路组织为 5 章，内容包括信息系统安全威胁、数据安全保护、身份认证与访问控制、网络安全保护和信息系统安全管理。

本书深入浅出，结构新颖，紧扣本质，适合教学，可以激发读者的求知欲。书中还配有丰富的实验和习题，供读者验证和自测。本书适合作为计算机科学与技术专业、信息管理与信息系统专业、网络专业和信息安全专业的“信息系统安全概论”课程的教材或教学参考书，也可供有关技术人员参考。



**图书在版编目（CIP）数据**

信息系统安全教程/张基温编著. —3 版. —北京：清华大学出版社，2017
（高等教育质量工程信息技术系列示范教材）
ISBN 978-7-302-48130-0

Ⅰ. ①信…　Ⅱ. ①张…　Ⅲ. ①信息系统－安全技术－高等学校－教材　Ⅳ. ①TP309

中国版本图书馆 CIP 数据核字（2017）第 208438 号

**责任编辑：** 白立军
**封面设计：** 常雪影
**责任校对：** 时翠兰
**责任印制：** 刘海龙

**出版发行：** 清华大学出版社
**网　　址：** http://www.tup.com.cn, http://www.wqbook.com
**地　　址：** 北京清华大学学研大厦 A 座　　**邮　　编：** 100084
**社 总 机：** 010-62770175　　**邮　　购：** 010-62786544
**投稿与读者服务：** 010-62776969，c-service@tup.tsinghua.edu.cn
**质 量 反 馈：** 010-62772015，zhiliang@tup.tsinghua.edu.cn
**课 件 下 载：** http://www.tup.com.cn,010-62795954
**印 装 者：** 北京鑫海金澳胶印有限公司
**经　　销：** 全国新华书店
**开　　本：** 185mm×260mm　　**印　　张：** 20.5　　**字　　数：** 482 千字
**版　　次：** 2007 年 8 月第 1 版　　2017 年 11 月第 3 版　　**印　　次：** 2017 年 11 月第 1 次印刷
**印　　数：** 1～2000
**定　　价：** 49.00 元

---

产品编号：075146-01

# 前　　言

## （一）

信息系统是重要的，重要的系统需要特别保护；计算机信息系统是复杂的，复杂的系统是脆弱的，脆弱的系统也需要特别保护；计算机信息系统是虚拟的，虚拟的系统给安全保护带来很大困难；现代信息系统是开放的，开放的系统会带来更多的风险。重要、复杂、虚拟和开放，也给人们带来研究的乐趣和商业机会。现在信息系统安全技术和产品已经大量涌现，并且还在不断发展。

本书主要介绍计算机信息系统的安全原理。作为一本原理类的教材，关键的问题是要梳理成合理而又容易理解和掌握的体系。在教学实践中，笔者反复探索，将安全理论梳理成如下 3 大类。

（1）攻防技术。恶意程序、网络攻击（黑客）、隔离（逻辑隔离——防火墙、物理隔离和电磁防护）、安全监控（IDS、网络诱骗和审计）、紧急响应和取证。

（2）安全信任体系。加密与信息隐藏、认证、安全协议。

（3）安全体系结构和评估标准。

这样的梳理基本上囊括了几乎所有的安全技术。

在内容的安排上，考虑了如下原则。

（1）尽量把与其他技术或概念相关的内容排在后面，即与其他内容相关最少的排在最前面。

（2）将能引起兴趣的内容排在前面，以使读者能有成就感。

（3）把安全体系结构和安全等级放在最后，这样不仅使其内容容易理解，而且也是对前面内容的总结和提高。

在内容的取舍上采取的原则是：重点内容详细介绍，次要内容只做一般介绍。

本书每章最后都配备了较多的习题，这些习题有如下不同的类型。

（1）有些要思考、总结。

（2）有些要进一步理解。

（3）有些要自己想象。

（4）有些要查找资料。

（5）有些要动手实验。

本书的第 1 版正是基于这些考虑而形成的。

## （二）

常言道：“道高一尺，魔高一丈。”信息系统安全是针对入侵和攻击采取的一系列安全策略和技术。然而，按照木桶原理，当一块短的木板被加长后，另一块次短的木板就变成

最短的木板了；而一种攻击被防御之后，新的攻击又会出现。在这个充满竞争的世界里，攻击与防御相伴而生并且永远不会完结，攻与防的博弈将在竞争中永无止境。一般说来，防御往往要比攻击付出更多的代价，其原因如下。

（1）攻击可以择机发起，防御必须随时警惕。

（2）攻击可以选择一个薄弱点实施，防御必须全线设防。

（3）攻击一般使用一种或几种技术，防御则需要考虑已知的所有技术。

（4）攻击可以肆意进行，防御必须遵循一定的规则。

社会总在发展，技术总在进步，攻击与防御也在博弈中相互促进。信息系统安全作为一门新兴的技术和学科，在本书第 1 版出版后的近 7 年间，又有了长足的发展。在读者和出版社的不断呼吁下，笔者下大力气进行全面修订，并在内容上进行了增删，删除了已经不再使用的技术，如 DES 加密算法等；增添了新技术的内容，如 AES 密码、流密码、僵尸网络等，使本书第 2 版于 2015 年出版。

教材不是一般的科技书，在编写过程中首先要考虑如何教的问题。为了更适应教学，第 2 版按照“威胁（第 1 章）—防护（第 2～4 章）—管理（第 5 章）”的体系进行组织。这相当于“提出问题—解决问题—总结提高”的思路。经过本人一个学期的试讲，效果不错。

本书第 3 版在第 2 版的基础上增加了近两年内引起人们注意的内容。

## （三）

本书的修订永远是进行时。在本书第 3 版即将出版之际，衷心地希望读者和有关专家能不吝指正，以便适当的时候再进一步修订。

在本书（包括第 1 版和第 2 版）的编写过程中，栾英姿、赵忠孝、张秋菊、董兆军、张展为、张友明、史林娟、张展赫、戴璐以及我的研究生陶利民、蒋中云、王玉斐、魏士婧、董瑜参与了部分编写工作，在此谨表谢意。

本书在编写过程中还参考了大量文献资料。这些资料有的引自国内外论文，有的引自其他著作，有的引自网站。虽本人尽心在参考文献中予以列出，但仍会有许多疏漏，同时也受篇幅所限，未能将所有参考资料一一列出。在此谨向有关作者致谢。

张基温

2017 年 6 月

# 目 录

**第 1 章 信息系统安全威胁**……1
1.1 恶意代码攻击……1
1.1.1 病毒……2
1.1.2 蠕虫……15
1.1.3 特洛伊木马……19
实验 1 判断并清除木马……27
1.1.4 陷门与黑客攻击……28
1.1.5 其他恶意代码……32
1.1.6 信息系统安全卫士……33
1.2 窃听型攻击……36
1.2.1 世界著名监听案例……36
1.2.2 声波窃听……41
1.2.3 电磁波窃听……44
1.2.4 光缆窃听……45
1.2.5 共享网络中的窃听……47
1.2.6 手机监听……48
1.2.7 NFC 泄露……51
1.3 系统扫描型攻击……51
1.3.1 网络扫描……52
1.3.2 漏洞扫描……59
实验 2 系统扫描……62
1.3.3 口令破解……63
1.4 欺骗型攻击……66
1.4.1 ARP 欺骗——交换网络监听……66
实验 3 监听器工具的使用……69
1.4.2 IP 源地址欺骗……70
1.4.3 路由欺骗……72
1.4.4 TCP 会话劫持……73
1.4.5 DNS 欺骗……74
1.4.6 Web 欺骗与钓鱼网站……76
1.4.7 伪基站攻击……80
1.5 数据驱动型攻击……82

1.5.1 缓冲区溢出攻击 ……82
1.5.2 格式化字符串攻击 ……84
1.6 拒绝服务攻击 ……86
1.6.1 拒绝服务攻击及其基本方法 ……86
1.6.2 分布式拒绝服务攻击 ……89
实验 4 拒绝服务攻击演示 ……93
1.6.3 僵尸网络 ……95
1.7 信息系统风险与安全策略 ……99
1.7.1 风险=脆弱性+威胁 ……99
1.7.2 信息系统安全策略 ……102
1.7.3 信息系统安全防御原则 ……103
习题 ……105

**第 2 章 数据安全保护** ……111
2.1 数据的机密性保护 ……111
2.1.1 数据加密基础 ……111
实验 5 加密博弈 ……114
2.1.2 数据加密体制 ……114
2.1.3 AES 算法 ……116
2.1.4 公开密钥算法 RSA ……121
2.1.5 密钥管理 ……123
2.1.6 流密码 ……128
2.1.7 量子加密 ……130
2.1.8 信息隐藏 ……133
2.2 消息认证——完整性保护 ……134
2.2.1 数据完整性保护与消息认证 ……134
2.2.2 MAC 函数 ……136
2.2.3 哈希函数 ……138
实验 6 实现报文认证算法 ……141
2.3 数字签名 ……141
2.3.1 数字签名及其特征 ……141
2.3.2 直接数字签名 ……142
2.3.3 有仲裁的数字签名 ……143
2.3.4 数字签名标准 DSA ……144
2.3.5 认证协议实例——SET ……145
实验 7 加密软件 PGP 的使用 ……149
习题 ……151

第 3 章　身份认证与访问控制……155
3.1　基于凭证比对的身份认证……155
3.1.1　生物特征身份认证……155
3.1.2　静态口令……157
3.1.3　动态口令……160
3.2　基于密钥分发的身份认证……162
3.2.1　公钥加密认证协议……162
3.2.2　单钥加密认证协议……163
3.2.3　Kerberos 认证系统……165
3.3　基于数字证书的身份认证……169
3.3.1　数字证书……169
3.3.2　X.509 证书标准……170
3.3.3　公开密钥基础设施……174
实验 8　证书制作及 CA 系统配置……175
3.4　信息系统访问授权……176
3.4.1　访问控制的二元关系描述……176
3.4.2　自主访问控制与强制访问控制……179
3.4.3　基于角色的访问控制策略……181
实验 9　用户账户管理与访问权限设置……181
习题……187

第 4 章　网络安全防护……190
4.1　网络防火墙……190
4.1.1　网络防火墙概述……190
4.1.2　防火墙技术之一——网络地址转换……193
4.1.3　防火墙技术之二——代理服务……196
4.1.4　防火墙技术之三——包过滤……199
4.1.5　防火墙技术之四——状态检测……204
4.1.6　网络防火墙部署……207
4.2　网络的物理隔离技术……210
4.2.1　物理隔离的概念……210
4.2.2　网络安全隔离卡……213
4.2.3　隔离集线器技术……214
4.2.4　网闸……215
4.3　Internet 安全协议……219
4.3.1　IPSec……219
4.3.2　SSL……223
4.3.3　VPN……227

实验 10　实现一个 VPN 连接 ······ 228
4.4　入侵检测系统 ······ 229
4.4.1　入侵检测及其模型 ······ 229
4.4.2　信息收集与数据分析 ······ 231
4.4.3　响应与报警策略 ······ 236
4.4.4　入侵检测器的部署与设置 ······ 237
4.5　网络诱骗 ······ 239
4.5.1　蜜罐主机技术 ······ 239
4.5.2　蜜网技术 ······ 240
4.5.3　常见网络诱骗工具及产品 ······ 241
习题 ······ 242

**第 5 章　信息系统安全管理 ······ 246**
5.1　信息系统应急响应 ······ 246
5.1.1　应急响应组织 ······ 246
5.1.2　信息系统安全保护制度 ······ 247
5.1.3　信息系统应急预案 ······ 249
5.1.4　灾难恢复 ······ 250
5.1.5　信息系统应急演练 ······ 254
5.2　数据备份、数据容错与数据容灾 ······ 256
5.2.1　数据备份 ······ 256
5.2.2　数据容错技术 ······ 261
5.2.3　数据容灾系统 ······ 263
5.3　数字证据获取 ······ 267
5.3.1　数字证据的特点与数字取证的基本原则 ······ 268
5.3.2　数字取证的一般步骤 ······ 269
5.3.3　数字取证的基本技术和工具 ······ 271
5.3.4　数字证据的法律问题 ······ 273
5.3.5　日志 ······ 274
5.4　信息系统安全风险评估与审计 ······ 277
5.4.1　信息系统安全风险评估及其目的 ······ 277
5.4.2　信息系统安全风险评估的准则与模式 ······ 278
5.4.3　信息系统安全风险评估过程 ······ 280
5.4.4　信息系统渗透测试 ······ 285
5.4.5　信息系统安全审计 ······ 289
5.5　信息系统安全测评准则 ······ 292
5.5.1　国际信息安全测评准则 ······ 292
5.5.2　中国信息系统安全保护等级划分准则 ······ 296

5.5.3 信息安全测评认证体系 ······299
5.6 开放系统互连安全体系结构 ······300
5.6.1 开放系统互连安全体系结构概述 ······301
5.6.2 OSI 安全体系结构的安全服务 ······302
5.6.3 OSI 七层中的安全服务配置 ······303
5.6.4 OSI 安全体系结构的安全机制 ······304
5.6.5 OSI 安全体系的安全管理 ······307
习题 ······310

**参考文献** ······313

# 第 1 章　信息系统安全威胁

信息系统安全威胁（thread）是指对信息系统的组成要素及其功能造成某种损害的潜在可能。信息系统是一种复杂的系统，复杂系统会有很多的脆弱之处。这是信息系统安全威胁的内因。信息系统也是现代社会的重要系统，重要系统往往是敌对方攻击的首选目标；并且信息系统所具有的神秘性往往会刺激好奇者、好胜者和恶作剧者的攻击兴趣。这是信息系统安全威胁的外因。

从技术上看，信息系统安全威胁涉及如下几个方面:

（1）机密性（confidentiality）威胁，即信息（数据）被非法泄露给那些未经授权掌握这一信息的实体，或被未经授权者窃取;

（2）完整性（integrity）威胁，即信息（数据）被未经授权的篡改;

（3）否认性（repudiation）威胁，也称为抵赖性（denial）威胁，即抵赖或否认已经完成的操作或承诺。

这些威胁有内部漏洞，也有外部攻击。本章从外部攻击的角度讨论信息系统的威胁。从形式上看，大致有如下 3 种攻击:

（1）恶意代码攻击，包括病毒、特洛伊木马、蠕虫、细菌、陷门和逻辑炸弹等;

（2）窃听攻击，包括声波窃听、电磁波窃听、光缆监听、手机窃听和网络窃听;

（3）黑客攻击，包括消息采集攻击、代码漏洞攻击、欺骗和会话劫持攻击、分布式攻击等。

## 1.1　恶意代码攻击

一般说来，恶意代码（malicious code）或恶意程序（malicious program）是指一类特殊的程序代码，它们通常在用户不知晓也未经授权的情况下潜入到计算机系统中，对系统产生不良影响。

在信息系统安全的反复博弈中，恶意代码不断花样翻新，种类不断增加。其中，最多见的是病毒（virus）、蠕虫（worm）、特洛伊木马（Trojan horse）、细菌、陷门（trap doors）、僵尸（bot）等。它们的区别主要是存在形式、自繁殖性、传播机制、运行机制和攻击机制等方面。表 1.1 为几种典型恶意代码的主要区别。

**表 1.1　几种典型恶意代码的主要区别**

| 恶意代码 | 存在形式 | 自繁殖性 | 传播机制 | 运行机制 | 攻击机制 |
|---|---|---|---|---|---|
| 病毒 | 寄生 | 有 | 文件感染传播 | 自主运行，条件触发 | 文件感染 |
| 蠕虫 | 独立 | 有 | 利用网络，主动传播 | 自主运行 | 消耗资源加恶意行为 |
| 特洛伊木马 | 欺骗性，独立 | 无 | 被植入 | 受控运行 | 窃取网络信息 |
| 细菌 | 独立 | 有 | 自传播 | 自主运行 | 消耗计算机资源 |

### 1.1.1 病毒

#### 1. 病毒的概念

在生物学界，病毒是一类没有细胞结构，但有遗传、复制等生命特征，主要由核酸和蛋白质组成的有机体。计算机病毒(computer virus)是指与生物界中的病毒有极为相似特征的程序。在《中华人民共和国计算机信息系统安全保护条例》中，明确定义为："计算机病毒，是指编制或者在计算机程序中插入的破坏计算机功能或者毁坏数据、影响计算机使用，并能自我复制的一组计算机指令或者程序代码。"

通常，人们也简单地把计算机病毒定义为：利用计算机软件与硬件的缺陷，破坏计算机数据并影响计算机正常工作的一组指令集或程序代码。更广义地说，凡是能够引起计算机故障，破坏计算机数据的程序代码都可称为计算机病毒。

#### 2. 病毒的特征

1）非授权执行性与寄生性

一个正常的程序是由用户调用的。被调用时，要从系统获得控制权，得到系统分配的相应资源，来实现用户要求的任务。病毒虽然具有正常程序所具有的一切特性，但是其执行具有非授权性。为此，它必须寄生在合法程序——宿主程序中，当用户运行正常程序时，病毒伺机取得系统的控制权，先于正常程序执行，并对用户呈透明状态。因此，寄生是病毒的重要特征。

就目前出现的各种病毒来看，病毒宿主体有两种：一种是磁盘引导扇区，另一种是可执行文件（.exe 或.com）。因为不论是磁盘引导扇区还是可执行文件，它们都有获取执行权的可能，这样病毒寄生在它们的上面，就可以在一定条件下获得执行权，从而使病毒得以进入计算机系统，并处于激活状态，然后进行病毒的动态传播和破坏活动。

病毒的寄生方式有两种：一种是替代法，另一种是链接法。替代法是指病毒用自己的部分或全部指令代码替代磁盘引导扇区或文件中的全部或部分内容。链接法则是指病毒将自身代码作为正常程序的一部分与原有正常程序链接在一起，病毒链接的位置可能在正常程序的首部、尾部或中间。寄生在磁盘引导扇区的病毒一般采用替代法，而寄生在可执行文件中的病毒一般采用链接法。

对于寄生在磁盘引导扇区的病毒来说，病毒引导程序占有了原系统引导程序的位置，并把原系统引导程序搬移到一个特定的地方。这样系统一启动，病毒引导模块就会自动地装入内存并获得执行权，然后该引导程序负责将病毒代码的传染模块和发作模块装入内存的适当位置，并采取常驻内存技术以保证这两个模块不会被覆盖，接着对这两个模块设定某种激活方式，使之在适当的时候获得执行权。处理完这些工作后，病毒引导模块将系统引导模块装入内存，使系统在带毒状态下运行。对于寄生在可执行文件中的病毒来说，病毒一般通过修改原有可执行文件，使该文件一执行就先转入病毒引导模块。该引导模块也完成把病毒的其他两个模块驻留内存及初始化的工作，然后把执行权交给执行文件，使系统及执行文件在带毒的状态下运行。

2）传染性

传染是病毒最本质的特征之一，是病毒的再生机制。生物界的病毒可以从一个生物体传播到另一个生物体，计算机病毒也可以从一个程序、部件或系统传播到另一个程序、部件或系统。在单机环境下，病毒的传染基本途径是通过磁盘引导扇区、操作系统文件或应用文件进行传染；在网络中，病毒主要是通过电子邮件、Web 页面、APP 等特殊文件和数据共享方式进行传染。

一般将传染分为被动传染和主动传染。通过网络传播或文件复制，使病毒由一个载体被携带到另一个载体，称为被动传染。病毒处于激活状态下，满足传染条件时，病毒从一个载体自我复制到另一个载体，称为主动传染。

从传染的时间性上看，传染分为立即传染和伺机传染。病毒代码在被执行瞬间，抢在宿主程序执行前感染其他程序，称为立即传染。病毒驻留内存后，当满足传染条件时才感染其他程序，称为伺机传染。

3）潜伏性与隐蔽性

病毒一旦取得系统控制权，可以在极短的时间内传染大量程序。但是，被感染的程序并不是立即表现出异常，病毒会潜伏下来，等待时机发作。

病毒的潜伏性依赖于其隐蔽性。为了隐蔽，病毒通常非常短小，一般只有几百字节或上千字节，此外还寄生于正常的程序或磁盘较隐蔽的地方，也有个别以隐含文件形式存在，不经过代码分析很难被发觉。

4）可触发性

潜伏下来的病毒一般要在一定的条件下才被激活，发起攻击。病毒具有判断这个条件的功能。下面列举一些病毒的触发（激活）条件。

（1）日期/时间触发：病毒读取系统时钟，判断是否激活。例如，“黑色星期五”逢 13 日的星期五发作，CIH 1.2 于每年的 4 月 26 日发作，CIH 1.3 则在 6 月 26 日发作，CIH 1.4 的发作日期则为每个月的 26 日。

（2）计数器触发：病毒内部设定一个计数单元，对系统事件进行计数，判定是否激活。例如，2708 病毒当系统启动次数达到 32 次时被激活，发起对串、并口地址的攻击。

（3）键触发：当输入某些字符时触发（如 AIDS 病毒，在输入 A、I、D、S 时发作）或以击键次数（如 Devil’s Dance 病毒在用户第 2000 次击键时被触发）或按键组合等为激发条件（如 Invader 病毒在按下 Ctrl+Alt+Delete 键时发作）。

（4）启动触发：以系统的启动次数作为触发条件。例如，Anti Tei 和 Telecom 病毒当系统第 400 次启动时被激活。

（5）感染触发：以感染文件个数、感染序列、感染磁盘数或感染失败数作为触发条件。例如，Black Monday 病毒在运行第 240 个染毒程序时被激活；VHP2 病毒每感染 8 个文件就会触发系统热启动操作等。

（6）条件触发：将多种条件综合使用，作为病毒代码的触发条件。

5）破坏性

破坏性体现了病毒的杀伤能力。大多数病毒具有破坏性，并且其破坏方式总在花样翻新。常见的病毒破坏性有以下几个方面。

（1）占用或消耗 CPU 资源及内存空间，导致一些大型程序运行受阻，系统性能下降。

（2）干扰系统运行，例如不执行命令、干扰内部命令的执行、虚发报警信息、打不开文件、内部栈溢出、占用特殊数据区、时钟倒转、重启动、死机、文件无法存盘、文件存盘时丢失字节、内存减小、格式化硬盘等。

（3）攻击 CMOS。CMOS 是保存系统参数（如系统时钟、磁盘类型、内存容量等）的重要场所。有的病毒（如 CIH 病毒）可以通过改写 CMOS 参数破坏系统硬件的运行。

（4）攻击系统数据区。硬盘的主引导记录、分区引导扇区、FAT（文件分配表）、文件目录等是系统重要的数据，这些数据一旦受损，将造成相关文件的破坏。

（5）攻击文件。现在发现的病毒中，大多数是文件型病毒。这些病毒会使染毒文件的长度、文件存盘时间和日期发生变化。

（6）干扰外部设备运行，如封锁键盘、产生换字、抹掉缓存区字符、输入紊乱、使屏幕显示混乱以及干扰声响、干扰打印机等。

（7）破坏网络系统的正常运行，例如发送垃圾邮件、占用带宽，使网络拒绝服务等。

**3. 病毒分类**

按照不同的分类标准，病毒可以分为不同的类型，下面介绍几种常用的分类方法。

1）按照病毒的攻击目标分类

（1）DOS 病毒：攻击 DOS 系统。

（2）UNIX/Linux 病毒：攻击 UNIX 或 Linux 系统。

（3）Windows 病毒：攻击 Windows 系统，如 CIH 病毒。

（4）OS/2 病毒：攻击 OS/2 系统。

（5）Macintosh 病毒：攻击 Macintosh 系统，如 Mac.simpsons 病毒。

（6）手机病毒。

（7）网络病毒。

2）按照病毒的寄生位置分类

（1）引导型病毒。引导型病毒是寄生在磁盘引导区的病毒。图 1.1 显示了硬盘的逻辑结构。可以看出，磁盘有两种引导区：主引导区和分区的引导区。因此也就有两种引导型病毒：主引导区病毒和分区引导病毒。

① 主引导区（Main Boot Record，MBR）病毒寄生在硬盘主引导程序所占据的硬盘 0 头 0 柱面第 1 个扇区中，典型的病毒有大麻病毒、2708 病毒、火炬病毒等。

② 分区引导（Boot Record，BR）病毒寄生在硬盘活动分区的逻辑 0 扇区（即 0 面 0 道第 1 个扇区），典型的病毒有 Brain、小球病毒、Girl 病毒等。

（2）文件型病毒。文件型病毒寄生在文件中，按照所寄生的文件类型可以分为 4 类。

① 可执行文件病毒，即寄生在扩展名为.com、.exe、.pe、.bat、.sys、.ovl 等文件的病

毒。一旦运行这类病毒的载体程序，就会使病毒注入、安装并驻留在内存中，伺机进行感染。感染了该类病毒的程序往往会减慢执行速度，甚至无法执行。

② 文档文件病毒或数据文件病毒，即寄生在 Word 文档、Excel 文档、Access 数据库文件等中的病毒。宏病毒（Macro）就是感染这些文件的病毒。

③ Web 文档病毒，如寄生在 HTML 文档和 HTM 文档的病毒。已经发现的 Web 病毒有 HTML/Prepend 和 HTML/ Redirect 等。

④ 目录文件病毒，如 DIR2 病毒。

（3）引导兼文件型病毒。这类病毒在感染文件时还伺机感染引导区，例如 CANCER 病毒、HAMMER Ⅴ病毒等。

（4）CMOS 病毒。CMOS 是保存系统参数和配置的重要地方，也存在一些没有使用的空间。CMOS 病毒就隐藏在这一空间中，从而可以躲避磁盘的格式化清除。

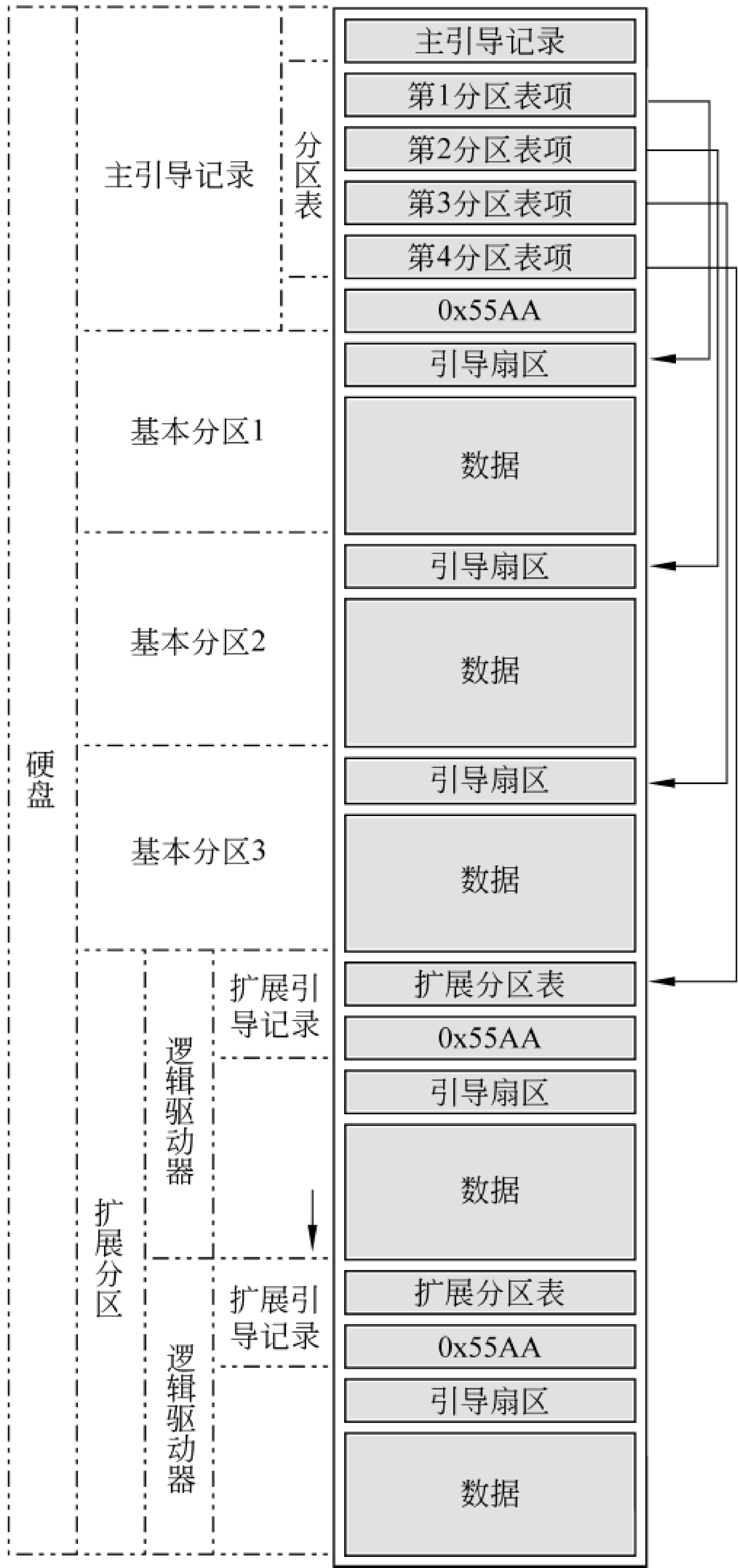



图 1.1　硬盘的逻辑结构

3）按照病毒是否驻留内存分类

（1）非驻留（nonresident）病毒。非驻留病毒选择磁盘上一个或多个文件，不等它们装入内存，就直接进行感染。

（2）驻留（resident）病毒。驻留病毒装入内存后，发现另一个系统运行的程序文件后进行传染。驻留病毒又可进一步分为以下几种。

① 高端驻留型。

② 常规驻留型。

③ 内存控制链驻留型。

④ 设备程序补丁驻留型。

4）按照病毒的表现形态分类

（1）多态病毒。这种病毒形态多样。它们在复制之前会不断改变形态以及自己的特征码，以躲避检测。例如，臭名昭著的“红色代码”病毒几乎每天变换一种形态。

（2）隐身病毒。隐身病毒对所隐身之处进行修改，以便藏身。隐身病毒分为两种情形。

① 规模修改：病毒隐藏感染一个程序之后，立即修改程序的规模。

② 读修改：病毒可以截获已感染引导区记录或文件的读请求并进行修改，以便隐藏。

（3）逆录病毒。这是一种攻击病毒查防软件的病毒。逆录病毒有 3 种攻击方式。

① 关闭病毒查防软件。

② 绕过病毒查防软件。

③ 破坏完整性校验软件中的完整性数据库。

（4）外壳病毒。这种病毒为自己添加一层保护外套，躲过病毒查防软件的检测、跟踪和拆卸。

（5）伴随病毒。这种病毒首先创建可执行文件，并在此基础上扩展，以便抢先执行。

（6）噬菌体病毒。这种病毒用自己的代码替代可执行代码，可以破坏接触到的任何可执行程序。

5）按照病毒的感染方式分类

按照感染方式，文件型病毒可以分为如图 1.2 所示的几种类型。

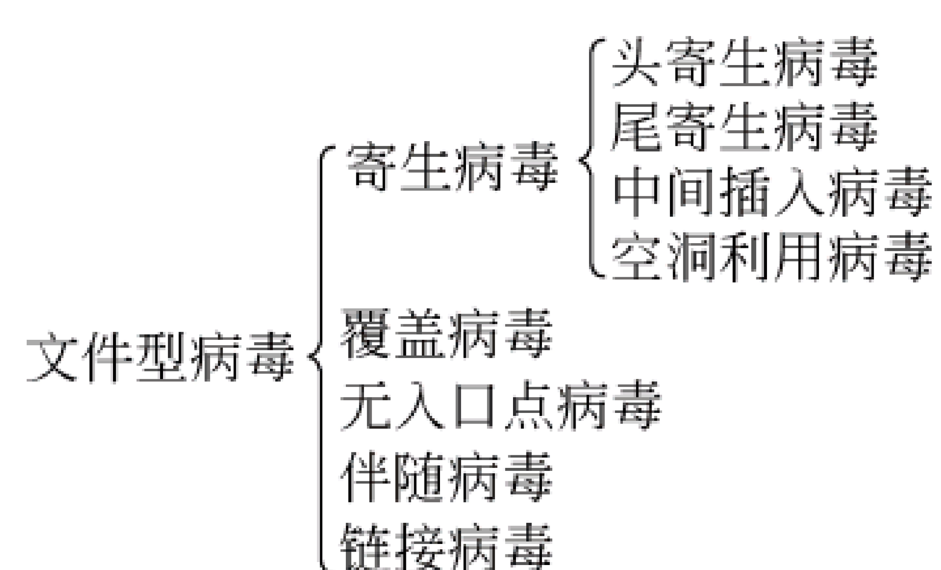



图 1.2　文件型病毒分类

（1）寄生病毒。这类病毒在感染的时候，将病毒代码加入正常程序之中，原来程序的功能部分或者全部被保留。根据病毒代码加入的方式不同，寄生病毒可以分为头寄生、尾寄生、中间插入和空洞利用 4 种。

头寄生是将病毒代码加入文件的头部。具体有两种方法：一种是将原来程序的前面一部分复制到程序的最后，空出病毒代码空间；另外一种是生成一个新的文件，首先在头的位置写上病毒代码，然后将原来的可执行文件放在病毒代码的后面。头寄生方式适合于不需要重新定位的文件，如批处理病毒和.com 文件。

尾寄生是将病毒代码加入文件的尾部，避开了文件重定位的问题，但为了先于宿主文件执行，需要修改文件头，使用跳转指令使病毒代码先执行。不过，修改头部也是一项复杂的工作。

中间插入是病毒将自己插入被感染的程序中，可以整段插入，也可以分成很多段，靠跳转指令连接。有的病毒通过压缩原来的代码的方法保持被感染文件的大小不变。

空洞利用多用于视窗环境下的可执行文件。因为视窗程序的结构复杂，其中都会有很多没有使用的部分，一般是空的段或者每个段的最后部分。病毒寻找这些没有使用的部分，然后将病毒代码分散到其中，从而实现了难以察觉的感染（CIH 病毒就使用了这种方法）。

（2）覆盖病毒。编制这种病毒的手法极其简单，是初期的病毒感染技术，它仅仅是直接用病毒代码替换被感染程序，使被感染的文件头变成病毒代码的文件头，不用做任何调整。

（3）无入口点病毒。这种病毒并不是真正没有入口点，只是在被感染程序执行的时候，并不立刻跳转到病毒的代码处开始执行，病毒代码无声无息地潜伏在被感染的程序中，可能在非常偶然的条件下才会被触发，开始执行。采用这种方式感染的病毒非常隐蔽，杀毒软件很难发现在程序的某个随机的部位有这样一些在程序运行过程中会被执行的病毒代码。

大量的可执行文件是使用 C 语言编写的，这些程序有这样一个特点，程序中会使用一

些基本的库函数，比如字符串处理、基本的输入输出等。为了使用这些库函数，编译器会在启动用户开发的程序之前增加一些代码对库进行初始化。这给了病毒一个机会，病毒可以寻找特定的初始化代码，并修改这段代码的开始语句，使得执行完病毒之后再执行通常的初始化工作。“纽克瑞希尔”病毒就采用了这种方法进行感染。

（4）伴随病毒。这种病毒不改变被感染的文件，而是为被感染的文件创建一个伴随文件（病毒文件），这样当执行被感染文件的时候，实际上执行的是病毒文件。

（5）链接病毒。这类病毒将自己隐藏在文件系统的某个地方，并使目录区中文件的开始簇指向病毒代码。这种感染方式的特点是每一个逻辑驱动器上只有一份病毒的副本。

6）按照病毒的破坏程度分类

按照破坏程度可将病毒分为以下几种类型。

（1）无害型：除了传染时减少磁盘的可用空间外，对系统没有其他影响。

（2）无危险型：这类病毒仅仅是减少内存、显示古怪图像、发出古怪声音等。

（3）危险型：这类病毒在计算机系统操作中造成严重的错误。

（4）非常危险型：这类病毒删除程序，破坏数据，清除系统内存区和操作系统中重要的信息。

### 4. 病毒状态

病毒一般会有如下 4 种状态。

（1）潜伏。病毒处于休眠状态，用户感觉不到病毒的存在。不过，有些病毒也可能没有潜伏期。

（2）感染。病毒感染其他程序。一般感染需要一定的条件。

（3）触发。病毒被某个条件激活，系统开始为其分配资源。

（4）发作。病毒开始运行，对系统形成一些破坏。

### 5. 病毒的逻辑结构

为了实现上述 4 种存在状态间的变换，病毒程序需要有如图 1.3 所示的 3 个模块：引导模块、感染模块和表现（破坏）模块。

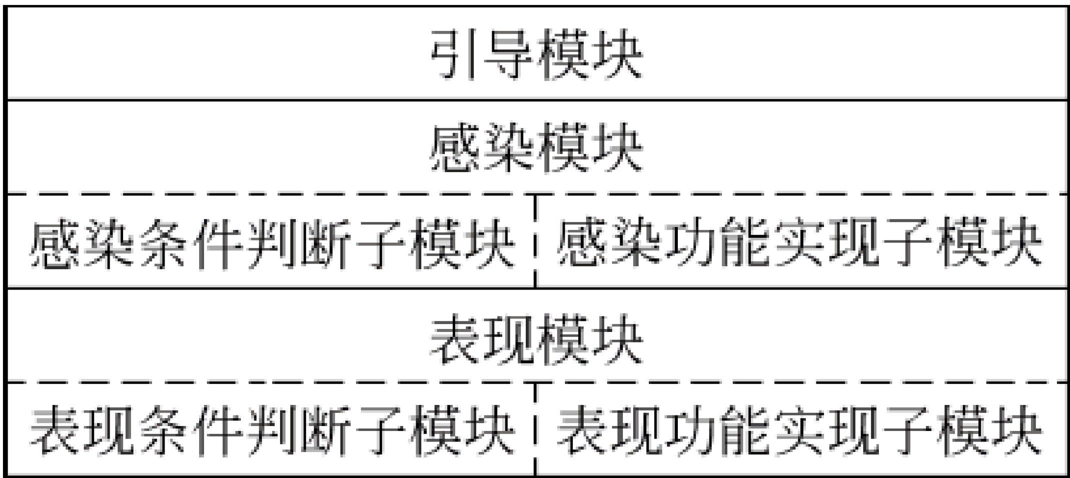



图 1.3　病毒的基本逻辑结构

1）引导模块

（1）引导模块的基本功能。引导模块也称为主控模块，主要实现如下功能。

① 将病毒装入内存，使感染和破坏（表现）模块处于活动的状态。

② 保护内存中的病毒代码不被覆盖。

③ 设置病毒的触发条件。

（2）引导过程如下。

① 检测运行环境，如操作系统类型、内存容量、现行区段、磁盘设置和显示器类型等。

② 驻留内存。自身的程序代码引入并驻留在内存中。

③ 窃取控制权。取代或扩充系统原有功能，并窃取系统的控制权，设置病毒的激活条件和触发条件，使病毒处于可激活状态，为感染模块做准备（如驻留内存、修改中断、修改高端内存、保存原中断向量等操作）。

④ 恢复系统功能。引导过程完成之后，病毒为了隐藏自己，等待时机进行感染和破坏，还要把控制权交还给系统。

（3）病毒代码引导模块的算法示例。

```
BootingModule(){
    将病毒代码寄生于宿主程序中;
    启动自我保护功能;
    设置感染条件;
    设置表现激活条件;
    宿主程序加载时，将病毒代码加载到内存;
}
```

2）感染模块

（1）病毒感染标志。一个程序感染了病毒后，往往会携带一个 ASCII 码形式的数字或字符串，称为病毒感染标志或病毒签名。不同的病毒，其感染标志的位置和内容不同。

具有感染标志的病毒在进行感染时，一般要查看感染对象是否已经带有感染标志，若有，就不再实施感染操作。杀毒软件也常常将感染标志作为病毒的特征码之一。人们也会利用感染标志实现病毒免疫。不过，病毒制造者也会利用感染标志实施欺骗，并且有一些病毒没有感染标志。

（2）感染模块的功能。感染机制的作用是在特定的感染条件下将病毒代码复制到被感染的目标上去。其主要功能包括 3 个方面。

① 寻找感染目标。

② 测试感染条件是否满足。

③ 实施感染。

（3）感染模块的组成。感染模块由以下两个子模块组成。

① 感染条件判断子模块。依据引导模块设置的感染条件，判断当前系统环境是否满足感染条件。

② 感染功能实现子模块。当一个感染条件满足时，启动感染功能，将病毒代码寄生在其他宿主程序上。

（4）病毒代码感染模块的算法示例。

```
InfectingModule(){
    实现感染目标程序的功能;
}
```

3）表现模块

（1）表现模块结构。表现机制的作用是在被感染系统上表现出特定现象，主要是产生破坏被感染系统的行为。大部分病毒都是在一定条件下才会被触发而发作。表现模块分为

两个子模块。

① 表现条件判断子模块。依据引导模块设置的感染条件，判断当前系统环境是否满足表现触发条件。

② 表现功能实现子模块。当一个表现条件满足时，启动病毒代码的表现功能，产生预定的效果。

（2）病毒代码表现模块的算法示例。

```
BehavingModule(){
        实现病毒代码的表现功能;
}
```

4）病毒代码的主函数算法

```
int main(){
      BootingModule();
      while(1){
            寻找感染目标;
            if(感染条件不满足) continue;
            InfectingModule();
            if(表现激活条件不满足) continue;
            BehavingModule();
            if(病毒代码需要退出)exit();
      }
}
```

**6. 引导型病毒解析**

如图 1.1 所示，硬盘可以分为多个分区，其中有一个主引导扇区，每个分区都有一个引导扇区。计算机开始工作时，首先执行的是引导程序。因此，引导扇区先于其他程序获得对 CPU 的控制。操作系统对于引导程序的辨认，不是依据内容，而是依据地址，即引导程序的入口地址是放在引导扇区的某个固定地方。这一特点给了引导型病毒一个机会：它可以偷梁换柱，将自己的入口地址放到原来的引导程序入口地址处。这样，当系统要装载引导程序时，实际是把一个引导型病毒装载到内存中。为了隐蔽自己，该病毒程序在自己进入内存后，再把引导程序装载到内存中。这样，病毒程序就可以驻留内存，监视系统运行，伺机传染和破坏。下面进一步介绍引导型病毒的工作原理。

1）引导型病毒的自举

引导型病毒为了能在操作系统被装入前先将自己装入，会先把 BOOT 引导程序搬移到另外一个地方，把自己的引导程序放在磁盘 0 面 0 道 1 扇区；同时修改 INT 13H 中断服务处理程序的入口地址，使之指向病毒引导模块。这样，当系统启动 ROM BIOS 之后，依托 BIOS 中断服务程序，就会先于 BOOT 执行病毒引导程序，把病毒代码装入内存，监视系统的运行，伺机感染插入的其他磁盘。图 1.4（a）简要描述了正常 DOS 自举过程，图 1.4（b）为带有病毒的 DOS 自举过程。

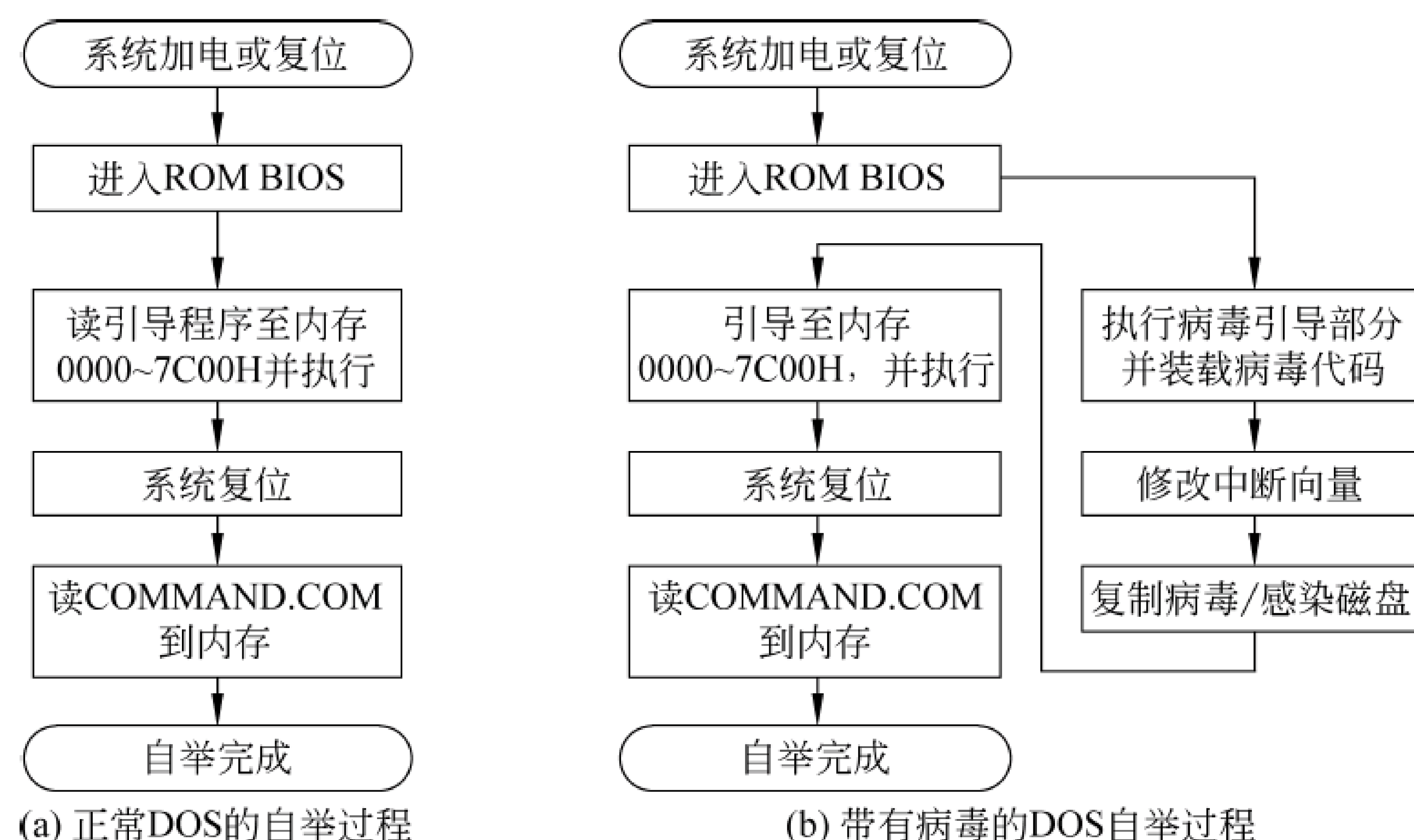



图 1.4　引导型病毒的一次活动过程

2）装入内存过程

（1）系统开机后，进入系统检测，检测正常后，从 0 面 0 道 1 扇区（即逻辑 0 扇区）读取信息到内存的 0000~7C00 处。

正常时，逻辑 0 扇区存放的是 BOOT 引导程序；操作系统感染了引导扇区病毒时，逻辑 0 扇区存放的是病毒引导部分，BOOT 引导程序被放到其他地方。例如，大麻病毒在软盘中将原 DOS 引导扇区搬移到 0 道 1 面 3 扇区，在硬盘中将原 DOS 引导扇区搬移到 0 道 0 面 7 扇区；香港病毒则将原 DOS 引导扇区搬移到 39 磁道第 8 扇区；Michelangelo 病毒在高密度软盘上是第 27 扇区，在硬盘上是 0 道 0 面 7 扇区。

（2）系统开始运行病毒引导部分，将病毒的其他部分读入内存的某一安全区，常驻内存，监视系统的运行。

（3）病毒修改 INT 13H 中断服务处理程序的入口地址，使之指向病毒控制模块并执行，以便必要时接管磁盘操作系统的控制权。BIOS INT 13H 调用是 BIOS 提供的磁盘基本输入输出中断调用，可以完成磁盘（包括硬盘和软盘）的复位、读写、校验、定位、诊断、格式化等操作。它采用 CHS 寻址（按柱面、磁头、扇区 3 个参数寻址，这是老的寻址方式。新的寻址方式是 LAB，即线性寻址），最大访问能力为 8GB 左右。

（4）病毒代码全部读入后，接着读入正常 BOOT 内容到内存 0000~7C00H 处，进行正常的启动过程（这时病毒代码已经全部被读入内存，不再需要病毒的引导部分）。

（5）病毒代码伺机等待随时感染新的系统盘或非系统盘。

3）攻击过程

病毒代码发现有可攻击的对象后，要进行下列工作。

（1）将目标盘的引导扇区读入内存，判断它是否感染了病毒。

（2）满足感染条件时，将病毒的全部或一部分写入引导区，把正常的磁盘引导区程序写入磁盘特定位置。

（3）返回正常的 INT 13H 中断服务处理程序，完成对目标盘的感染过程。

**7. Win32 PE 文件病毒解析**

1）Win32 PE 文件格式

Win32 PE 文件就是 Win32（Windows 95/98/2000/XP）环境下的 PE 格式的可执行文件。为了了解病毒对它的感染机理，首先介绍 PE 文件的结构和运行机制。PE 文件的格式如图 1.5 所示。

PE（Portable Executable）即可移植的执行体。所有 PE 文件必须以一个具有重定位功能的可执行文件格式 DOS MZ（MZ 是主要作者 Mark Zbikowski 名字的缩写）头开始。MZ 头中包括各种说明数据，如第一句可执行代码执行指令时所需要的文件入口点、堆栈的位置、重定位表等，操作系统根据文件头的信息将代码部分装入内存，然后根据重定位表修正代码，并在设置好堆栈后从文件头中指定的入口开始执行。因此 DOS 可以把程序放在任何它想要的地方。图 1.6 是 MZ 格式的可执行文件的简单结构示意图。

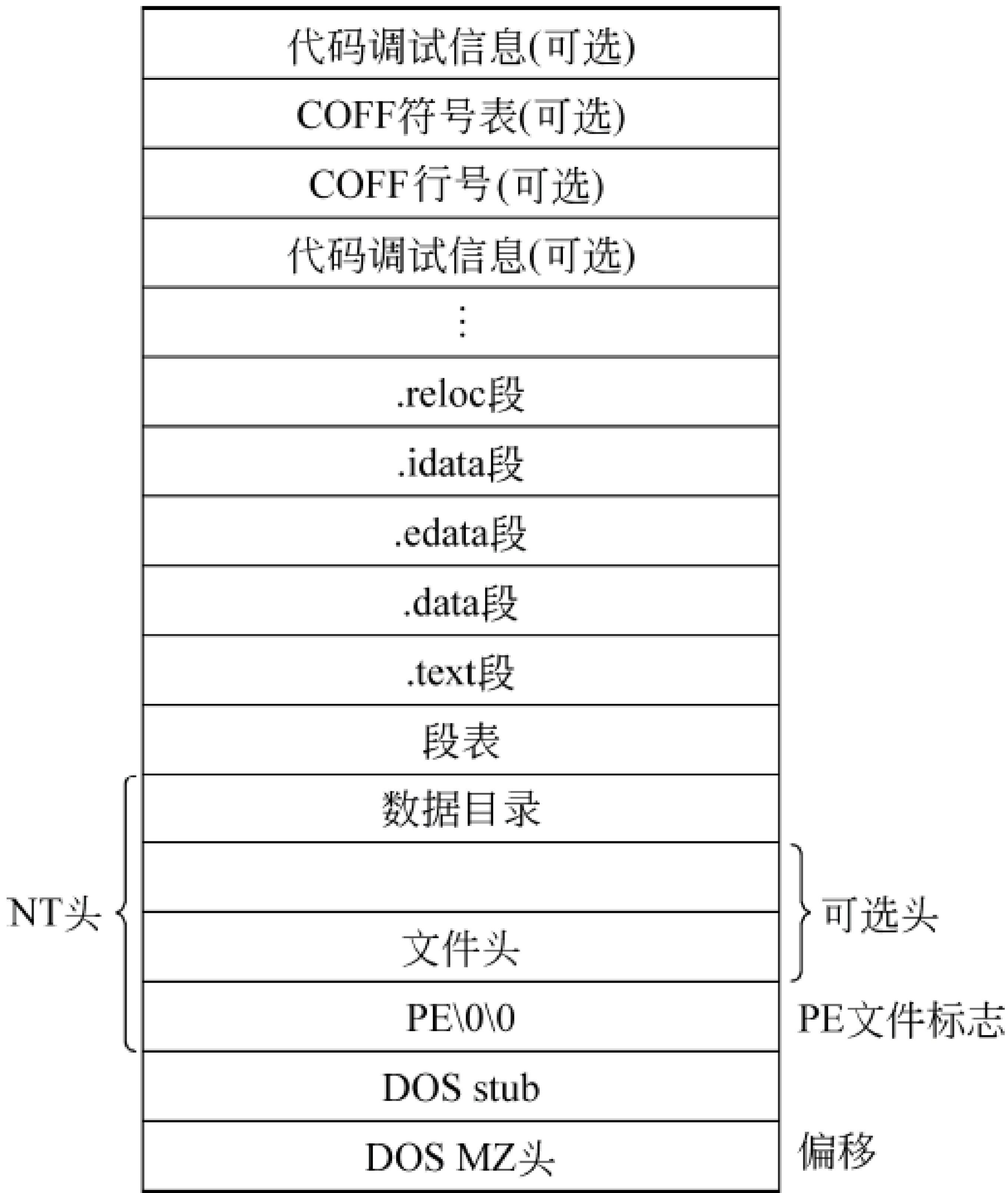



图 1.5　PE 文件的格式

| | |
|---|---|
| MZ 标志 | MZ 文件头 |
| 其他信息 | |
| 重定位表的字节偏移量 | |
| 重定位表 | 重定位表 |
| 可重定位程序映像 | 二进制代码 |

图 1.6　MZ 格式的可执行文件的简单结构

DOS stub 是一个极小（几百个字节）的 DOS 程序，用于输出警告，如“该程序不能在 DOS 模式下运行”等。当 Win32 把一个 PE 文件映像加载到内存时，内存映像文件的第一个字节对应到 DOS stub 的第一个字节。

PE 头长度为 1024B，是 PE 文件的标志。执行体在支持 PE 文件的操作系统中执行时，PE 装载器将从 DOS MZ 头中找到 PE 头的偏移量。

PE 文件的内容部分由一些称为段的块组成。每段是一块具有共同属性的数据。段数写在段表中。

2）PE 文件的装载过程

（1）PE 文件被执行时，PE 装载器检查 DOS MZ 头中的 PE 头偏移量，找到了，就跳转

到 PE 头。

（2）PE 检查器检查 PE 头的有效性。若有效，就跳转到 PE 头的尾部。

（3）读取段表信息，通过文件映射，将段映射到内存，同时附上段表中指定的段属性。

（4）PE 文件映射到内存后，PE 装载器处理 PE 文件中的有关逻辑。

3）重定位

对于正常的程序来说，数据的内存存储位置在编译时就已经计算好了，程序可按照这个地址直接装入。而病毒代码可能依附在宿主程序的不同位置，当病毒随着宿主程序装载到内存后，病毒中数据的位置也会随之发生变化。由于指令是通过地址引用数据的，地址的不准确将导致病毒代码的不正确执行。为此，有必要对病毒代码中的数据进行重定位。

4）获取 API 函数地址

在 Win32 环境中，系统功能调用不是通过中断实现，而是通过调用 API 函数实现。因此，获取 API 函数的入口地址非常重要。但是，Win32 PE 病毒与普通的 Win32 PE 程序不同。普通的 Win32 PE 程序里有一个引入函数表，程序通过这个表可以找到代码段中所用的 API 函数在动态链接库中的真实地址。调用 API 函数时，可以通过该引入函数表找到相应 API 函数的真正执行地址。但是，Win32 PE 病毒只有一个代码段，并不存在引入函数表，因此不能直接用真实地址调用 API 函数。获取 API 地址是病毒的一个重要技术。

5）其他机制

（1）搜索目标文件。通常通过 API 函数 FindFirstFile 和 FindNextFile 进行。

（2）内存文件映射。使用内存文件映射进行文件读写。

（3）感染其他文件。

（4）返回到宿主程序。

**8. 计算机病毒防范**

1）强化安全管理，加强防范体系与规范建设

没有一个完善的防范体系，一切防范措施都将滞后于病毒的危害。病毒防范制度是防范体系中每个主体都必须遵守的行为规程。没有制度，防范体系就不可能很好地运作，就不可能达到预期的效果。必须依照防范体系对防范制度的要求，结合实际情况，建立符合自身特点的防范制度。

为了统筹全国的病毒防治工作，2000 年 5 月在原计算机病毒防治产品检验中心的基础上成立了国家计算机病毒应急处理中心。国家计算机病毒应急处理中心的工作任务是：充分调动国内防治病毒的力量，快速发现病毒疫情，快速做出反应，快速处置，及时消除病毒，防止病毒对我国的计算机网络和信息系统造成重大的破坏，确保我国信息产业安全健康发展。

为使中国计算机病毒防治工作走上法制轨道，国家于 1994 年 2 月颁布了《中华人民共和国计算机信息系统安全保护条例》，在此基础上又颁布了《计算机病毒防治管理办法》，还在新修订的《中华人民共和国刑法》中对故意制造、传播病毒的行为规定了相应的处罚

办法。

在一个具体部门中，也要建立自己的行之有效的计算机病毒防治制度，并严格执行。一般说来，应当做到以下几方面。

（1）系统管理员的口令应严格管理，不使之泄露，要不定期地予以更换，保护网络系统不被非法存取，不被感染上病毒或遭受破坏。

（2）应用程序软件的安装，应由系统管理员进行或由系统管理员临时授权进行，以保障网络用户使用共享资源时总是安全无毒的。

（3）系统管理员对网络内的共享电子邮件系统、共享存储区域和用户卷应定期进行病毒扫描，发现异常情况及时处理。条件许可时，还应在应用程序卷中安装最新版本的杀病毒软件供用户使用。

（4）网络系统管理员在做好日常管理事务的同时，还要拟订应急措施，及时发现病毒感染迹象。一旦出现病毒传播迹象，应立即隔离被感染的计算机系统和网络，并进行处理，而不应当带毒继续工作。要按照特别情况清查整个网络，切断病毒传播的途径，保障正常工作的进行。

（5）对新购置的计算机硬软件系统进行测试。

（6）单台计算机系统的安全使用要采取以下措施。

① 当在一台计算机中使用在其他计算机中用过的移动存储器时，应当先对其进行病毒检测。

② 对重点保护的计算机系统应做到专机、专盘、专人、专用。

（7）封闭的使用环境中是不会自然产生病毒代码的。对于网络计算机系统，除了要首先保证自己使用的计算机安全外，还应采取下列针对网络的防杀病毒措施。

① 安装网络服务器时，应保证安装环境和网络操作系统本身没有感染病毒。

② 安装网络服务器时，应将文件系统划分成多个文件卷系统。一旦系统卷受到某种损伤，导致服务器瘫痪，就可以通过重装系统卷恢复网络操作系统，使服务器能马上投入运行，而装在共享的应用程序卷和用户卷内的程序和数据文件不会受到任何损伤。如果用户卷内由于病毒或使用上的原因导致存储空间拥塞时，系统卷不会受影响，不会导致网络系统运行失常。这种划分还十分有利于系统管理员设置网络安全存取权限，保证网络系统不受病毒感染和破坏。

③ 要用硬盘启动网络服务器，否则在受到引导型病毒感染和破坏后，遭受损失的将不仅仅是一台个人计算机，而会影响整个网络的中枢。

④ 在网络服务器上必须安装真正有效的防杀病毒软件，并经常进行升级。必要的时候还可以在网关、路由器上安装病毒防火墙产品，从网络出入口保护整个网络不受病毒的侵害。

⑤ 不随便直接运行或直接打开电子函件中夹带的附件文件，不随意下载软件，尤其是一些可执行文件和 Office 文档；即使下载了，也要先用最新的防杀病毒软件进行检查。

（8）重要数据文件要有备份。例如：

① 硬盘分区表、引导扇区等的关键数据应作备份并妥善保管，以便在进行系统维护和修复工作时作为参考。

② 重要数据文件定期进行备份工作，不要等到被病毒破坏、计算机硬件或软件出现故障使用户数据受到损伤时再去急救。

2）计算机病毒检测

病毒是一段程序代码，即使它隐藏得很好，也会留下许多痕迹。通过对这些蛛丝马迹的判别，发现病毒的特征和名称，就称为查毒。目前使用的查毒方法有以下几种。

（1）现象观测法。现象观测法是根据病毒代码发作前、发作时和发作后的表现现象，推断发现病毒代码。

（2）进程监视法。进程监视会观察到系统的活动状况，同时也会拦截所有可疑行为。例如，多数个人计算机的 BIOS 都有防病毒设置，当这些设置打开时，就允许计算机拦截所有对系统主引导记录进行写入的企图。

（3）比较法。比较法是用原始的或正常的文件与被检测的文件进行比较。按照比较的内容有以下几种方法。

① 长度（内容）比较法。

② 内存比较法。

③ 中断比较法。

④ 校验和比较法。

比较法可以通过感染实验室法进行。它先运行一些确切知道不带毒的正常程序，然后观察这些正常程序的长度和校验和，如果发现有的程序增长，或者校验和有变化，就可以断定系统中有病毒。

（4）特征码法。特征码法是将所有病毒的病毒特征码加以剖析，把分析得到的这些病毒独有的特征搜集在一个病毒特征码资料库（病毒库）中。检测时，以扫描的方式将待检测程序与病毒库中的病毒特征码一一对比；发现有相同的代码，则可判定该程序已遭病毒感染。这种方法是许多病毒检测工具的基础。但是，这种方法检测不出未知病毒。

（5）软件模拟法。软件模拟法是一种软件分析器，用软件方法来模拟和分析程序的运行。这种方法后来演变为虚拟机上进行的查毒、启发式查毒等技术，是相对成熟的技术。

（6）分析法。分析法是适用于反病毒技术人员的病毒检测方法，分为静态分析和动态分析两种方法。

静态分析法是用 DEBUG 等反汇编程序，将病毒代码打印成反汇编后的程序清单进行分析，看病毒分成哪些模块，使用了哪些系统调用，采用了哪些技巧，如何将病毒感染文件的过程翻转为清除病毒、修复文件的过程，哪些代码可被用作特征代码以及如何防御这种病毒。

动态分析法是利用 DEBUG 等调试工具，在内存带病毒的情况下对病毒进行动态跟踪，观察病毒的具体工作过程，以进一步在静态分析的基础上理解病毒工作的原理。

在病毒编码比较简单的情况下，动态分析不是必需的。当病毒采用了较多的技术手段时，必须使用动、静相结合的分析方法才能完成整个分析过程。

### 1.1.2 蠕虫

#### 1. 蠕虫及其特征

1982 年，Xerox PARC 的 John F.Shoch 等人为了进行分布式计算的模型实验，编写了称为蠕虫（worm）的程序。可他们没有想到，这种“可以自我复制”并可以“从一台计算机移动到另一台计算机”的程序，后来竟给计算机界带来了巨大的灾难。1988 年被罗伯特·莫里斯（Robert Morris）释放的 Morris 蠕虫在 Internet 上爆发，在几个小时之内迅速感染了所能找到的存在漏洞的计算机。如图 1.7 所示，Internet 上的蠕虫袭击最初缓慢地增加，到 2000 年后开始呈指数上升。

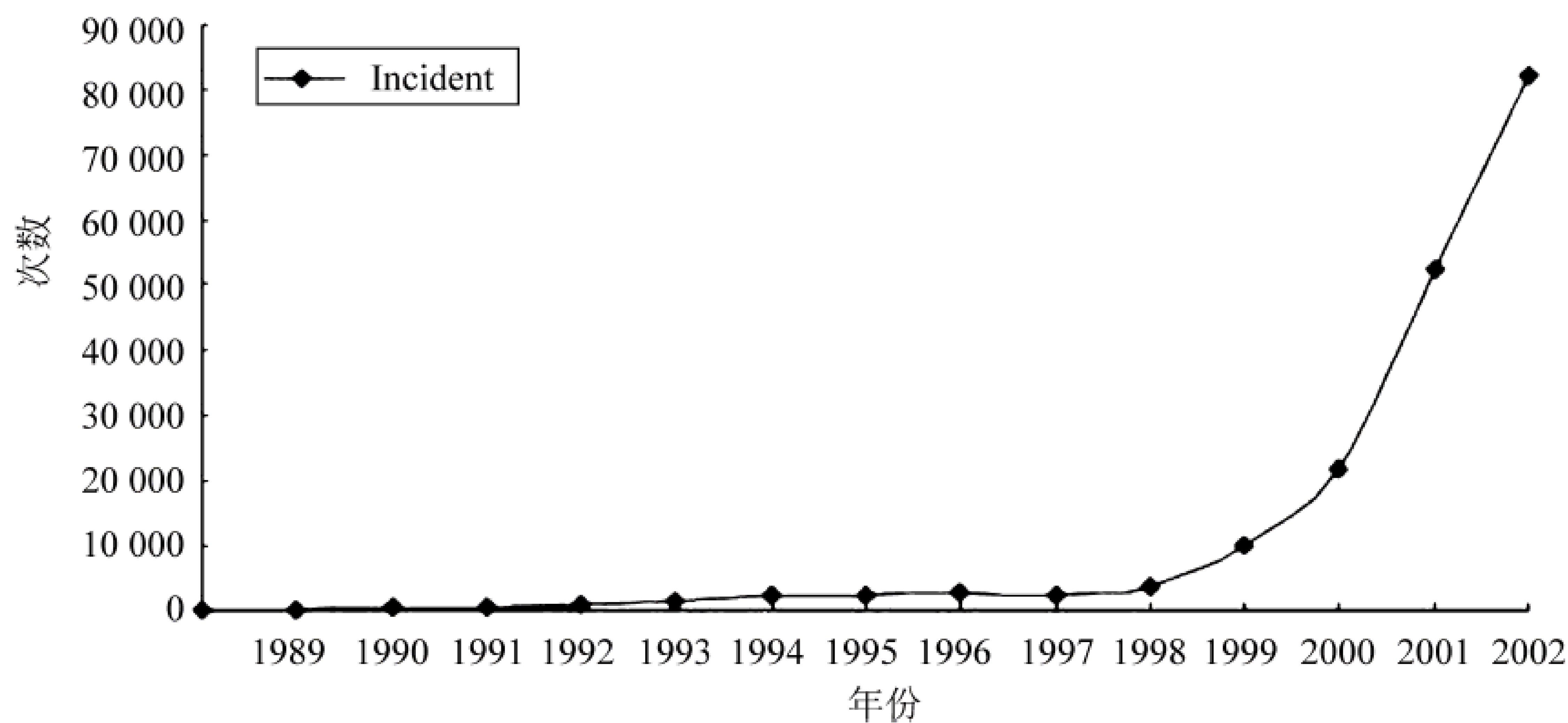



图 1.7　CERT（计算机紧急响应小组）统计的 1988—2002 年蠕虫事件的发生次数

下面进一步说明蠕虫的特点。

（1）存在的独立性。病毒具有寄生性，寄生在宿主文件中；而蠕虫是独立存在的程序个体。

（2）攻击的对象是计算机。病毒的攻击对象是文件系统，而蠕虫的攻击对象是计算机系统。

（3）传播的主动性。蠕虫与病毒都采用自传播方式。感染型病毒可以搜索并感染同一台计算机上能够访问到的其他文件。但与病毒不同的是，蠕虫传播的主动性非常强，它会不断地复制，并非常努力地通过各种途径将自身或变种传播到其他计算机系统中，因此可能造成更广泛的危害。蠕虫的传播途径很多，如操作系统漏洞、电子邮件、移动设备、即时通信等，也可以随网络攻击传播，使人防不胜防。

（4）感染的反复性。病毒与蠕虫都具有感染性，都可以自我复制。但是，病毒与蠕虫的感染机制有 4 点不同。

① 病毒感染是一个将病毒代码嵌入到宿主程序的过程，而蠕虫的感染是自身的复制。

② 病毒的感染目标针对本地程序（文件），而蠕虫是针对网络上的其他计算机。

③ 病毒是在宿主程序运行时被触发进行感染，而蠕虫是通过系统漏洞进行感染。

④ 蠕虫是一种独立程序，也可以作为病毒的寄生体，携带病毒，并在发作时释放病毒，进行双重感染。

（5）行踪的隐蔽性。由于蠕虫传播过程的主动性，蠕虫不需要像病毒那样由计算机使用者的操作触发，并且还常常伪装成有用的软件，因而难以察觉。例如，2011 年 9 月 5 日，首个 QQ 群蠕虫（Pincav）被截获。该蠕虫伪装成电视棒破解程序欺骗网民下载，盗取魔兽、邮箱及社交网络账号，计算机被感染后，蠕虫会自动访问 QQ 群共享空间来进行传播。其感染量每天约 2 万个。到了 2013—2014 年间，QQ 群蠕虫升级为第四代，它们利用大众猎奇心理，将其程序改名为“××视频助手.exe”或“××视频偷看神器.exe”进行伪装，来吸引网民点击。如果网民信以为真，双击运行，蠕虫就会劫持网民的 QQ，把推广消息转发到 QQ 群共享和空间说说，甚至发送病毒邮件给好友。该蠕虫的最终目的是在中毒计算机上安装一大堆流氓软件以牟取暴利；在 Android 手机上则会弹出全屏广告，并在后台偷偷下载色情应用，严重影响手机的正常使用。

（6）破坏的严重性。病毒虽然对系统性能有影响，但破坏的主要是文件系统。而蠕虫主要是利用系统及网络漏洞影响系统和网络性能，降低系统性能。例如，它们的快速复制以及在传播过程中的大面积漏洞搜索，会造成巨量的数据流量，导致网络拥塞甚至瘫痪；对一般系统来说，多个副本形成大量进程，会大量耗费系统资源，导致系统性能下降，对网络服务器尤为明显。

非但如此，蠕虫通常还会在传播的同时执行一些其他的恶意行为，以达到自己的目的。例如，2007 年 1 月流行的“熊猫烧香”蠕虫病毒，利用微软公司视窗操作系统的漏洞进行传播。计算机感染这一病毒后，会不断自动拨号上网，并利用文件中的地址信息或者网络共享进行传播，最终破坏用户的大部分重要数据。

蠕虫破坏的严重性造成了巨大的经济损失。例如：

1988 年 11 月 2 日，Morris 蠕虫发作，一夜之间攻击了约 6200 台 VAX 系列小型机和 Sun 工作站。Purdue 大学 Gene Spafford 估计整个经济损失为 20 万美元，而美国病毒代码协会的 John McAfee 的报告认定的损失大约为 9600 万美元。

1998 年爆发的 CIH 蠕虫在世界范围内造成 2000 万～8000 万美元的损失。

1999 年“美丽杀手”蠕虫使政府部门和一些大公司紧急关闭了网络服务器，经济损失超过 12 亿美元。

2000 年 5 月“爱虫”开始流传，造成大量计算机感染，迄今造成的损失超过 100 亿美元。

2001 年 7 月 19 日，“红色代码”（Code Red）蠕虫爆发，几个小时内就攻击了 25 万台计算机，造成的损失超过 1 亿美元。之后该蠕虫产生了威力更强的几个变种，在世界范围内造成大约 280 万美元的损失。

2001 年 12 月开始流传的“求职信”造成大量邮件服务器堵塞，损失达数百亿美元。

2003 年 1 月，Sql 蠕虫王造成网络大面积瘫痪，银行自动提款机运行中断，直接经济损失超过 26 亿美元。

2003 年 1 月 25 日，Slammer 首次出现，其目标是服务器，在 10 分钟内感染了 7.5 万台计算机，曾使整个韩国的计算机网络瘫痪了 12 小时，全球受感染的服务器超过了 50 万台，5 天之内造成的损失超过 10 亿美元。

2003 年夏季，“冲击波”（Blaster）爆发，数十万台计算机被感染，给全球造成了 20 亿～

100 亿美元的损失。

2003 年 8 月 19 日，Sobig 的变种“霸王虫”爆发，在最初的 4 小时内自身复制了 100 万次，给全球带来 50 亿～100 亿美元的损失。

2004 年 1 月 18 日，“贝革热”（Bagle）爆发，给全球带来数千万美元的损失。

2004 年 1 月 26 日出现在网络上的 MyDoom，传播速度大大超过了“霸王虫”。“霸王虫”在传播高峰期的记录是每 17 封邮件就有 1 封被感染，而 MyDoom 在 1 月 28 日就创下了每 12 封邮件中就有 1 封被感染的记录。至 1 月 30 日，感染率在大客户中是每 10 封邮件中就有 1 封被感染，在小客户中是每 3 封邮件中就有 1 封被感染。在 2004 年最烧钱的病毒代码评比中，MyDoom 称冠。

2004 年 4 月 30 日，“震荡波”（Sasser）爆发，给全球带来数千万美元的损失。

2005 年 8 月 16 日，Zotob 蠕虫及其数个变种流行。

2006 年 6 月 2 日，“维金”蠕虫（Viking）被截获，截至该年年底，受攻击用户数量达 1 740 679。

2006 年 10 月 16 日，“熊猫烧香”（别名“尼姆亚”或“武汉男生”又化身为“金猪报喜”）爆发，据中国国家计算机网络应急处理中心估计，其造成的损失超过 76 亿元。

2008 年 11 月，“扫荡波”蠕虫被发现。当时，它已经造成大量企业用户局域网瘫痪，数十万用户网络崩溃。

2008 年年底，Conficker 蠕虫开始传播。一旦计算机被感染，就被加入一个大规模的僵尸网络，并被蠕虫作者控制。在首次被检测到之后，Conficker 已经感染了数百万台计算机和多个国家的企业网络。

2008 年年底，“刻毒虫”（Kido）开始流行，到 2010 年已经成为感染面积最大的蠕虫。

2010 年 6 月，“震网”（Stuxnet）首次被白俄罗斯安全公司 VirusBlokAda 发现。实际上它的传播从 2009 年 6 月开始，甚至更早。它是首个针对工业控制系统的蠕虫，也是已知的第一个以关键工业基础设施为目标的蠕虫。

2017 年 5 月 12 日，一种名为 WannaCry（又称为 Wanna Decryptor）的蠕虫勒索病毒在全球范围大爆发。当用户主机系统被该勒索软件入侵后，即会将用户主机上的重要文件，如照片、图片、文档、压缩包、音频、视频、可执行程序等几乎所有类型的文件进行加密，被加密的文件后缀名被统一修改为 WNCRY。同时弹出一个勒索对话框，要求支付价值相当于 300 美元的比特币才可解锁。否则用户只能通过重装操作系统的方式来解除勒索行为，而用户的重要数据文件不能直接恢复。WannaCry 软件大小为 3.3MB，由不法分子利用 NSA（National Security Agency，美国国家安全局）泄露的危险漏洞 EternalBlue（永恒之蓝）进行传播。勒索病毒肆虐，俨然是一场全球性互联网灾难。据当时的不完全统计，在短短一周内，至少 150 个国家、30 万名用户中招，造成损失达 80 亿美元。事发后，各国立即采取行动，国际刑警组织、欧洲警察署、美国联邦调查局（FBI）、英国国家犯罪局等紧急启动调查，追踪犯罪者。2017 年 5 月 13 日晚间，由一名英国研究员于无意间发现的 WannaCry 隐藏开关（Kill Switch）域名，意外地遏制了该病毒的进一步大规模扩散。

从上述讨论可以看出，蠕虫虽然与病毒有些不同，但也有许多共同之处。如果将凡是能够引起计算机故障，破坏计算机数据的程序均统称为病毒代码，那么，从这个意义上说，蠕

虫也应当是一种病毒。它以计算机为载体，以网络为攻击对象，是通过网络传播的恶性病毒。

**2. 蠕虫代码的功能结构**

一个蠕虫程序的基本功能包括传播模块、隐藏模块和目的模块 3 部分。

1）传播模块

传播模块用于实现蠕虫的自动入侵功能。没有蠕虫的传播技术，也就谈不上蠕虫技术了。传播模块由扫描子模块、攻击子模块和复制子模块组成。

（1）扫描子模块负责探测存在漏洞的主机。当程序向某个主机发送探测漏洞的信息并收到成功的反馈信息后，就会得到一个可传播的对象。

一般说来，蠕虫希望隐蔽地传播，并尽快地传播到更多的主机。根据这一原则，扫描模块采取的扫描策略是：随机选取一段 IP 地址，然后对这一地址段上的主机进行扫描。

差的扫描程序并不知道一段地址是否已经被扫描过，只是随机地扫描 Internet，很有可能重复扫描一个地址段。于是，蠕虫传播得越广，网上的扫描包越多，即使探测包很小，但积少成多，就会引起严重的网络拥塞。

扫描策略改进的原则是，尽量减少重复的扫描，使扫描发送的数据包尽量少，并保证扫描覆盖尽量大的范围。按照这一原则，可以有如下一些策略。

① 在网段的选择上，可以主要对当前主机所在网段进行扫描，对外网段随机选择几个小的 IP 地址段进行扫描。

② 对扫描次数进行限制。

③ 将扫描分布在不同的时间段进行，不集中在某一时间内。

④ 针对不同的漏洞设计不同的探测包，提高扫描效率。例如，对远程缓冲区溢出漏洞，通过发出溢出代码进行探测。对 Web CGI 漏洞，发出一个特殊的 HTTP 请求探测。

（2）攻击子模块按照漏洞攻击步骤自动攻击已经找到的攻击对象，获得一个 shell，就拥有了对整个系统的控制权。对 Windows 2000 来说，就是 cmd.exe。

（3）复制子模块通过原主机和新主机的交互，将蠕虫程序复制到新主机并启动，实际上是一个文件传输过程。

图 1.8 表明了蠕虫的工作过程。蠕虫首先随机生成一个 IP 地址作为要攻击的对象，接着对被攻击的对象进行扫描，探测有无存在漏洞的主机。当程序向某个主机发送探测漏洞的信息并收到成功的反馈信息后，就得到一个可传播的对象，随后就可以将蠕虫主体迁移到目标主机。然后，蠕虫程序进入被感染的系统，对目标主机进行现场处理。现场处理部

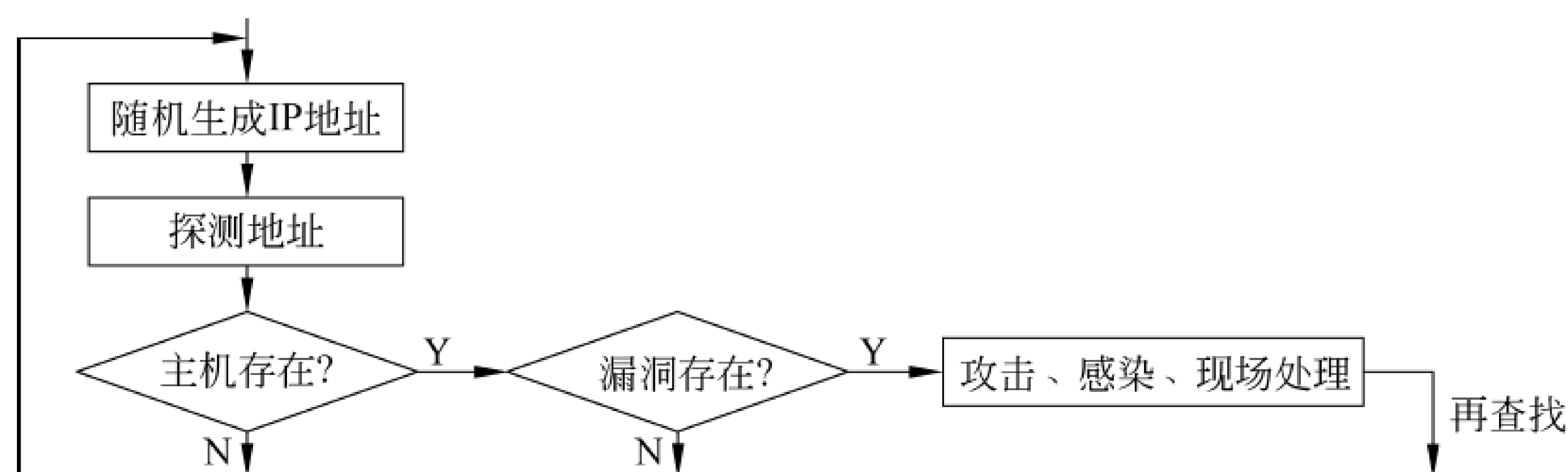



图 1.8　蠕虫的工作流程

分的工作包括隐藏、信息搜集等。蠕虫入侵计算机系统之后，会在被感染的计算机上产生自己的多个副本，每个副本启动搜索程序寻找新的攻击目标。一般要重复上述过程 $m$ 次（$m$ 为蠕虫产生的繁殖副本数量）。不同的蠕虫采取的 IP 生成策略可能并不相同，甚至随机生成。各个步骤的繁简程度也不同，有的十分复杂，有的则非常简单。

2）隐藏模块

隐藏模块负责在侵入主机后隐藏蠕虫程序，防止被用户发现。

蠕虫为了不被发现，就要采用一些隐藏技术。下面介绍蠕虫的几种隐藏手法。

（1）修改蠕虫在系统中的进程号和进程名称，掩盖蠕虫启动的时间记录。此方法在 Windows 95/98 下可以使用 RegisterServiceProcess API 函数使得进程不可见。但是，在 Windows NT/2000 下，由于没有这个函数，方法就要复杂一些了：只能在 psapi.dll 的 EnumProcess API 上设置“钩子”，建立一个虚的进程查看函数。

（2）将蠕虫复制到一个目录下，并更换文件名为已经运行的服务名称，使任务管理器不能终止蠕虫运行。这时要参考 ADV API32.DLL 中的 OpenSCManagerA 和 CreateServiceA API 函数。

（3）删除自己。

在 Windows 95 系统中，可以采用 DeleteFile API 函数。

在 Windows 98/NT/2000 中，只能在系统下次启动时删除自己。常用的方法是在注册表中加入如下一条：

```
HKLM\\SOFTWARE\\Microsoft\\Windows\\CurrentVersion\\RUNONCE%COMSPEC%/CDEL<PATH_TO_WORM\\
WORM_FILE_NAME.EXE>
```

应在尖括号括起的部分输入蠕虫的路径和文件名；然后重新启动操作系统。

3）目的模块

该模块实现对计算机的控制、监视或破坏等功能，其核心是攻击子模块。

### 1.1.3 特洛伊木马

#### 1. 特洛伊木马及其特征

古希腊诗人荷马（Homer）在其史诗《伊利亚特》（*The Iliad*）中描述了这样一个故事：希腊王的王妃海伦被特洛伊（Troy）的王子掠走，希腊王在攻打特洛伊城（位于今土耳其西北面的恰纳卡莱省的希沙利克）时，使用了木马计（the strategy of Trojan horse），在巨大的木马内装满了士兵，然后假装撤退，把木马留下。特洛伊人把木马当作战利品拉回特洛伊城内。到了夜间，木马内的士兵钻出来作为内应，打开城门，希腊王得以攻下特洛伊城。此后，人们就把特洛伊木马（Trojan horse）作为伪装的内部颠覆者的代名词。图 1.9 为重造的木马与特洛伊遗址。

RFC 1244（Request For Comments:1244）中，关于特洛伊木马程序的定义是：特洛伊木马（以下简称木马）程序是一种恶意程序，它能提供一些有用的或者令人感兴趣的功能；但是还具有用户不知道的其他功能，例如，在用户不知晓的情况下复制文件或窃取密码。



(a) 重造的木马



(b) 特洛伊遗址

图 1.9　重造的木马与特洛伊遗址

简单地说，凡是人们能在本地计算机上操作的功能，木马基本上都能实现。

进入 21 世纪后，木马已经成为恶意程序中增长较快的一种。金山毒霸全球病毒疫情监测系统的数据表明，多年来在每年的新增恶意程序中，木马一直占据 70%左右，表 1.2 为 2006—2009 年的数据。此外，木马的破坏性大大增强。2010 年的数据表明，新型木马的破坏性超过传统木马的 10 倍。

**表 1.2　2006—2009 年金山毒霸全球病毒疫情监测系统截获的新增病毒、木马数量**

| 年 份 | 新增病毒、木马总数量 | 新增木马数量 | 比例（%） |
|---|---|---|---|
| 2006 | 240 156 | | 73 |
| 2007 | 11 147 | 7 659 | 68.7 |
| 2008 | 13 899 717 | 7 801 911 | 56.13 |
| 2009 | 20 684 223 | 15 223 588 | 73.6 |

木马是一种危害性极大的恶意代码。它执行远程非法操作者的指令，进行数据和文件的窃取、篡改和破坏，释放病毒，以及使系统自毁等任务。下面介绍它的特征。

（1）目的性和功能特殊性。一般说来，每个木马程序都赋有特定的使命，其活动目的都比较清楚，例如盗号木马、网银木马、下载木马等。木马的功能都是十分特殊的，除了普通的文件操作以外，还有些木马具有搜索高速缓存中的口令、设置口令、扫描目标计算机的 IP 地址、进行键盘记录、远程注册表的操作以及锁定鼠标等功能。

（2）非授权性与受控性。非授权性是指木马的运行不需由受攻击系统用户授权，受控性是指木马的活动大都是由攻击者控制的。一旦控制端与服务器端建立连接后，控制端将窃取用户密码，获取大部分操作权限，如修改文件、修改注册表、重启或关闭服务器端操作系统、断开网络连接、控制服务器端鼠标和键盘、监视服务器端桌面操作、查看服务器端进程等。这些权限不是用户授权的，而是木马自己窃取的。

（3）非自繁殖性、非自传播性与预入性。一般说来，病毒具有极强的感染性，蠕虫具有很强大的传播性，而木马不具备繁殖性和自动感染的功能，其传播是通过一些手段植入的。例如，可以在系统软件和应用软件的文件传播中人为植入，也可以在系统或软件设计

时被故意放置进来。例如，微软公司曾在其操作系统设计时故意放置了一个木马程序，可以将客户的相关信息发回到其总部。

（4）欺骗性。隐藏是一切恶意代码的存在之本。而木马为了获得非授权的服务，还要通过欺骗进行隐藏。例如，它们使用的是常见的文件名或扩展名，如 dll\\win\\sys\\explorer 等字样；或者仿制一些不易被人区别的文件名，如字母 l 与数字 1、字母 o 与数字 0，木马经常修改基本文件中的这些难以分辨的字符，更有甚者干脆借用系统文件中已有的文件名，只不过将它保存在不同的路径之中。木马通过这些手段便可以隐藏自己，更重要的是，通过偷梁换柱的行动，让用户把它当作要运行的软件启动。有一类网购木马可以利用多款银行交易系统接口，后台自动查询银行卡余额，将中毒网民银行卡的所有余额一次窃走。例如，“秒余额”网购木马采用的骗术是：当网民在淘宝网买完东西，骗子说你的订单被卡单了，需要联系某某人处理。不明真相的网民联系后，会被诱导运行不明程序，这个程序就是网购木马。中毒后，只要网民继续购物，就会造成网银资金损失。

### 2. 特洛伊木马分类

1）根据攻击动作方式分类

（1）远程控制型。远程控制型是木马程序的主流。远程控制就是在计算机间通过某种协议（如 TCP/IP）建立一个数据通道。通道的一端发送命令，另一端解释并执行该命令，并通过该通道返回信息。简单地说，就是采用 Client/Server（客户机/服务器，简称 C/S）工作模式。

采用 C/S 模式的木马程序都由两部分组成：一部分为被控端（通常是监听端口的 Server 端），另一部分称为控制端（通常是主动发起连接的 Client 端）。被控端的主要任务是隐藏在被控主机的系统内部，并打开一个监听端口，就像隐藏在木马中的战士等待着攻击的时机，当接收到来自控制端的连接请求后，主线程立即创建一个子线程并把请求交给它处理，同时继续监听其他的请求。控制端的任务只是发送命令，并正确地接收返回信息。

这种类型的木马运行起来非常简单，只要先运行服务器端程序，同时获得远程主机的 IP 地址，控制者就能任意访问被控制端的计算机，从而使远程控制者在本地计算机上做任何想做的事情。

（2）信息窃取型。信息窃取型木马的目的是收集系统上的敏感信息，例如用户登录类型、用户名、口令和密码等。这种木马一般不需要客户端，运行时不会监听端口，只悄悄地在后台运行，一边收集敏感信息，一边不断检测系统的状态。一旦发现系统已经连接到 Internet 上，就在受害者不知情的情形下将收集的信息通过一些常用的传输方式（如电子邮件、ICQ、FTP）把它们发送到指定的地方。

（3）键盘记录型。键盘记录型木马只做一件事情，就是将受害者的键盘敲击完整地记录在 LOG 文件中。

（4）毁坏型。毁坏型木马以毁坏并删除文件（如受害者计算机上 DLL、INI 或 EXE 格式的文件）为主要目的。

2）根据木马程序的功能分类

（1）网络游戏木马。网络游戏木马常常以盗取网游账号密码为目的，它通常采用记录

用户键盘输入、Hook 游戏进程 API 函数等方法获取用户的密码和账号。窃取到的信息一般通过发送电子邮件或向远程脚本程序提交的方式发送给木马控制者。

（2）网银木马。网银木马针对网上交易系统，以盗取用户的卡号、密码甚至安全证书为目的，常常造成受害用户的惨重损失。这类木马控制者可能会首先对某银行的网上交易系统进行仔细分析，然后针对安全薄弱环节编写病毒程序。如 2004 年的“网银大盗”病毒，在用户进入该银行网银登录页面时，会自动把页面换成安全性能较差，但依然能够运转的老版页面，然后记录用户在此页面上填写的卡号和密码；“网银大盗 3”利用招行网银专业版的备份安全证书功能，可以盗取安全证书。

（3）即时通信软件木马。常见的即时通信类木马一般有如下 3 种。

① 发送消息型。这类木马能自动发送含有恶意网址的消息，让收到消息的用户点击网址中毒，用户中毒后又会向更多好友发送病毒消息。此类病毒的常用技术是搜索聊天窗口，进而控制该窗口自动发送文本内容。发送消息型木马常常充当网游木马的广告，如“武汉男生 2005”木马可以通过 MSN、QQ、微信等多种聊天软件发送带毒网址，其主要功能是盗取游戏的账号和密码。

② 盗号型。这类木马的工作原理和网游木马类似。病毒作者盗得他人账号后，可能偷窥聊天记录等隐私内容，或将账号卖掉。

③ 传播自身型。这类木马可以发布消息或文件并通过 QQ 聊天软件发送自身进行传播。具体方法是搜寻到聊天窗口后，对聊天窗口进行控制，以便发送文件或消息。

（4）网页点击类木马。网页点击类木马会恶意模拟用户点击广告等动作，在短时间内可以产生数以万计的点击量。木马作者的编写目的一般是为了赚取高额的广告推广费用。此类木马的技术简单，一般只是向服务器发送 HTTP GET 请求。

（5）下载类木马。这种木马程序的体积一般很小，其功能是从网络上下载其他病毒程序或安装广告软件。由于体积很小，它们更容易传播，传播速度也更快。通常功能强大、体积也很大的后门类病毒，如“灰鸽子”“黑洞”等，传播时都单独编写一个小巧的下载型木马，用户中毒后会把后门主程序下载到本机运行。

（6）代理类木马。用户感染代理类木马后，会在本机开启 HTTP、SOCKS 等代理服务功能。黑客把受感染计算机作为跳板，以被感染用户的身份进行活动，达到隐藏自己的目的。

### 3. 木马的功能与结构

1）木马的功能

一般说来，木马具有如下一些功能。

（1）远程监视、控制。远程监视和控制是木马最主要的功能，通过这个功能，可以让黑客就像使用自己的计算机一样使用被种植了木马的计算机。同时为了不引起对方的察觉，也可以只是远程监视，则对方的一举一动都在黑客的监视之下。并且当对方有摄像头时，还可以自动启动摄像头捕捉图像，相当于监视对方的环境。

（2）远程管理。远程管理包括很多功能，比如远程文件管理、远程 Telnet、远程注册表管理等，这些都是为了方便黑客控制主机而设置的。

（3）获得并发送主机消息。在客户端选择被控服务器端，单击“远程控制命令”标签，

弹出系统信息，这时可以得到被控服务器端的详细信息，然后将这些消息发送到客户端。

（4）修改系统注册表。木马可以使黑客单击客户端上的“注册表编辑器”标签，展开远程主机，并在远程主机的注册表上进行修改、添加、删除等一系列操作。

（5）执行远程命令。执行远程攻击者的命令。

2）木马软件的结构

木马软件一般由木马配置程序、木马控制程序和木马程序（服务器端程序和受控端程序）组成。

（1）木马配置程序。木马配置程序用于设置木马程序的端口号、触发条件和木马名称等，使其在服务器端隐藏更深。

（2）木马控制程序。木马控制程序用于控制远程木马服务器，给服务器发送指令，同时接收服务器传送来的数据。

（3）木马程序（服务器程序）。木马程序驻留在目标系统中，非法获取其操作权，负责接收控制指令，并根据指令或配置发送数据给控制端。

#### 4. 木马植入技术

为了有效地工作，木马一般都采用客户机/服务器形式，即由攻击者控制的客户端程序和运行在被控计算机端的服务器程序组成。木马的植入，就是将木马的服务器程序放置到目标主机上。下面介绍木马的几种植入形式。

（1）手工放置。手工放置是比较简单而常见的做法。手工放置分为本地放置和远程放置两种。本地放置就是直接在计算机上进行安装。远程放置就是通过常规攻击手段使获得目标主机的上传权限后，将木马上传到目标计算机上，然后通过其他方法使木马程序运行起来。

（2）以邮件附件的形式传播。控制端将木马改头换面，然后将木马程序添加到附件中，发送给收件人。

（3）通过 OICQ 对话，利用文件传送功能发送伪装了的木马程序。

（4）捆绑文件。这种伪装手段是将木马捆绑到一个安装程序上，当安装程序运行时，木马在用户毫无察觉的情况下偷偷地进入了系统。被捆绑的文件一般是可执行文件（即.exe、.com 一类的文件）。

（5）通过病毒或蠕虫程序传播。

（6）通过 U 盘或光盘传播。

#### 5. 木马隐藏技术

隐藏是一切恶意代码生存之本，欺骗是通过伪装来实现隐藏的一种技巧。

1）木马隐藏的一般策略

（1）隐蔽进程。服务器端想要隐藏木马，可以伪隐藏，也可以真隐藏。伪隐藏是指程序的进程仍然存在，只不过是让它消失在进程列表里。真隐藏则是让程序彻底消失，不以一个进程或者服务的方式工作。

伪隐藏的方法比较简单。在 Windows 9x 系统中，只要把木马服务器端的程序注册为一个服务（在后台工作的进程）就可以了。这样，木马程序就会从任务列表中消失，系统不再认为其是一个进程，当按下 Ctrl+Alt+Delete 键的时候，也就看不到这个进程。对于 Windows NT、Windows 2000 等，通过服务管理器，则要使用 API 的拦截技术，通过建立一个后台的系统钩子，拦截 PSAPI 的 EnumProcessModules 等相关的函数来实现对进程和服务的遍历调用的控制，当检测到进程 ID（PID）为木马程序的服务器端进程的时候直接跳过，这样就实现了进程的隐藏。

当进程为真隐藏的时候，木马完全融进了系统内核，因此就不把它做成一个应用程序，其服务器程序运行之后，就不具备一般进程的特征，也不具备服务的特征。

（2）修改文件标志或文件名，将木马文件伪装成图像、HTML、TXT、ZIP 等文件。一般来说，安装到系统文件夹中的木马的文件名是固定的，因此只要根据一些查杀木马的文章按图索骥，在系统文件夹查找特定的文件，就可以断定中了什么木马。所以现在有很多木马都允许控制端自由定制安装后的木马文件名，这样用户就很难判断所感染的木马类型了。

（3）定制端口。很多老式的木马端口都是固定的，这给用户判断是否感染了木马带来了方便，只要查一下特定的端口就知道感染了什么木马，所以现在很多新式的木马都加入了定制端口的功能，控制端用户可以在 1024～65 535 之间任选一个端口作为木马端口（一般不选 1024 以下的端口），这样就给用户判断系统所感染的木马类型带来了麻烦。

（4）伪装成应用程序扩展组件。将木马程序写成任何类型的文件（如 .dll、.ocx 等），然后挂在十分出名的软件中，因为人们一般不怀疑这些软件。

（5）错觉欺骗。利用人的错觉，例如故意混淆文件名中的 1（数字）与 l（L 的小写）、0（数字）与 o（字母）或 O（字母）。

（6）合并程序欺骗。合并程序就是将两个或多个可执行文件结合为一个文件，使这些可执行文件能同时执行。木马的合并欺骗就是将木马绑定到应用程序中。

（7）出错显示——施放烟雾弹。有一定木马知识的人都知道，如果打开一个文件，没有任何反应，这很可能就是一个木马程序，为了能蒙蔽用户，有的木马程序被点击时，会弹出一个错误提示框（这当然是假的），显示诸如“文件已破坏，无法打开！”之类的信息，当用户信以为真时，木马却悄悄侵入了系统。

2）便于木马启动的隐藏方式

植入目标主机的木马只有启动运行，才能开启后门为攻击者提供服务。为了便于启动，可以将木马程序隐藏在下列位置。

（1）集成到程序中。作为一种客户机/服务器程序，木马为了不让用户能轻易地把它删除，就常常集成到程序里，一旦用户激活木马程序，木马文件就会和某一应用程序捆绑在一起，然后上传到服务器端覆盖原文件，这样即使木马被删除了，只要运行捆绑了木马的应用程序，木马又会被安装上去了。如果它绑定到系统文件，那么每一次系统启动均会启动木马。

（2）隐藏在配置文件中。利用配置文件的特殊作用，木马很容易在计算机中运行、发作，从而偷窥或者监视其他计算机。不过，这种方式不是很隐蔽，容易被发现。

（3）潜伏在 Win.ini 中。Win.ini 通常是木马比较惬意的潜伏地方。因为 Win.ini 的\[windows\]字段中有启动命令 load=和 run=。在一般情况下=后面是空白的，这为木马程序留了一个合适的隐藏场所。

（4）隐藏在 System.ini 中。Windows 安装目录下的 System.ini 为木马提供了下列隐藏场所。

① 木马程序可以接在 System.ini 的[boot]字段的 hell=Explorer.exe 后面。

② System.ini 中的[386Enh]字段的“driver=路径\程序名”也有可能被木马所利用。

③ System.ini 中的[mic]、[drivers]、[drivers32]这 3 个字段也是起到加载驱动程序的作用，也是增添木马程序的场所。

（5）隐蔽在 Winstart.bat 中。Winstart.bat 也是一个能自动被 Windows 加载运行的文件，它多数情况下为应用程序及 Windows 自动生成，在执行了 Win. com 并加载了多数驱动程序之后开始执行(这一点可通过启动时按 F8 键再选择逐步跟踪启动过程的启动方式得知)。由于 autoexec.bat 的功能可以由 Winstart.bat 代替完成，因此木马完全可以像在 autoexec.bat 中那样被加载运行。

（6）捆绑在启动配置文件中。黑客利用应用程序的启动配置文件能启动程序的特点，将制作好的带有木马启动命令的同名文件上传到服务器端覆盖这同名文件，这样就可以达到启动木马的目的了。

（7）设置在超链接中。木马的主人在网页上放置恶意代码，引诱用户点击。

（8）加载程序到启动组。木马隐藏在启动组，虽然不是十分隐蔽，但非常便于自动加载运行。常见的启动组如下。

① “开始”菜单中的启动项，对应的文件夹是 C:\Documents and Settings\用户名\[开始]菜单\程序\启动。

② 注册表[HKEY_CURRENT_USER\SOFTWARE\Microsoft\Windows\CurrentVersion\Run]项。

③ 注册表[HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows\CurrentVersion\Run]项。

（9）注册成为服务项。将服务器端程序注册为一个自启动的服务也是木马常用的手段，其在注册表中的键值是[HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet \Services\]，比如“灰鸽子”就是这样。

**6. 木马的连接与远程控制**

1）木马的连接

从连接方法来分，木马可以分为 3 类：正向连接型、反向连接型和反弹连接型。

（1）正向连接型。这种类型的客户端连接服务器的时候是直接根据服务器的 IP 地址和端口来进行连接，比如 Penumbra。这种方法直观、简单，但同时也存在相应的缺陷。它要求知道对方的 IP。这对于不是固定 IP 的被控端而言，过一段时间 IP 改变后，控制端就无法连接了。为了弥补这个问题，一些正向连接型的木马工具提供了开机发邮件通知等方法，在被控端一开机时就把主机信息通知给控制端。但对于有防火墙的主机，这种方法就不一定能成功，毕竟防火墙对于这种直接连接是比较敏感的。基于以上缺陷，发展出了后面的

反向连接型木马。

（2）反向连接型。在反向连接型木马系统中，不再是由控制端去连接被控端，而是控制端自动监听，由被控端来进行连接，例如“流萤”。这个办法可以很好地解决正向连接所遇到的问题，但这种方法要求控制端有一个固定的公网 IP。如果控制端 IP 是自动分配的话，过一段时间后 IP 发生变化，则被控端就不可能连接到了。

（3）反弹连接型。这种木马由反向连接型发展而来，它也是由被控端去连接控制端，但其被控端的木马程序不知道控制端的 IP，而是知道一个固定的网页文件地址，这个地址一般是网上的免费空间。典型的例子就是“灰鸽子”。

2）木马的远程控制

木马连接建立后，控制端端口和木马端口之间将会出现一条通道。控制端上的控制端程序可通过这条通道与服务端上的木马程序取得联系，并通过木马程序对服务器端进行远程控制。

（1）窃取密码。一切以明文的形式、*形式或缓存在 Cache 中的密码都能被木马侦测到。此外很多木马还提供击键记录功能，它将会通过记录服务器端每次击键的位置和动作，计算出输入的内容。所以一旦有木马入侵，密码将很容易被窃取。

（2）文件操作。控制端可通过远程控制对服务器端上的文件进行删除、新建、修改、上传、下载、运行、更改属性等一系列操作，基本涵盖 Windows 平台上所有的文件操作功能。

（3）修改注册表。控制端可任意修改服务器端注册表，包括删除、新建，或修改主键、子键、键值。有了这项功能，控制端就可以禁止服务端 U 盘、光驱的使用，锁住服务器端的注册表，将服务器端上木马的触发条件设置得更隐蔽的一系列高级操作。

（4）系统操作。这项内容包括重启或关闭服务器端操作系统，断开服务器端网络连接，控制服务器端的鼠标和键盘，监视服务器端桌面操作，查看服务器端进程等，控制端甚至可以随时给服务器端发送信息。

3）木马的数据传送

木马程序的数据传递方法有很多种，通常是靠 TCP、UDP 传输数据。这时可以利用 Winsock 与目标机的指定端口建立起连接，使用 send 和 recv 等 API 进行数据的传递，但是这种方法的隐蔽性比较差，往往容易被一些工具软件查看到。例如，在命令行状态下使用 netstat 命令，就可以查看到当前活动的 TCP、UDP 连接。

4）数据传送时躲避侦察的方法

木马常用以下 3 种方法躲避数据传送时的侦察。

（1）合并端口法：使用特殊的手段，在一个端口上同时绑定两个 TCP 或者 UDP 连接（例如 80 端口的 HTTP），通过把自己的木马端口绑定于特定的服务端口（例如 80 端口的 HTTP）之上达到隐藏端口的目的。

（2）修改 ICMP（Internet Control Message Protocol）头法：使用 ICMP 进行数据发送，同时修改 ICMP 头，加入木马的控制字段。这样的木马具备很多新的特点，如不占用端口，

使用户难以发觉，并可以穿透一些防火墙，从而增大了防范的难度。

（3）为了避免被发现，木马程序必须很好地控制数据传输量，例如把屏幕画面切分为多个部分，并将画面存储为 JPG 格式，使压缩率变高，使数据变得十分小，甚至在屏幕没有改变的情况下传送的数据量为 0。

## 实验 1　判断并清除木马

### 1. 实验目的

（1）掌握在 TCP/IP 系统中判断木马的方法。

（2）掌握手工清除常见木马的基本方法。

### 2. 查看开放端口进行木马判断

当前最常见的木马是基于 TCP/IP，按照 C/S 模式工作的。这样，被种上木马的服务器就会打开监听端口等待连接。例如，冰河木马的监听端口是 7626，Back Orifice 2000 使用的端口是 54320，因此，通过查看本机开放端口，就可以判定自己的计算机是否中了木马或其他黑客程序。

下面介绍几种查看开放端口的方法。

（1）使用 Windows 自身带的 netstat 命令。netstat 命令是 Windows 自带的运行在 TCP/IP 环境的一个命令。它可以显示并统计有关连接和侦听端口。其命令格式如下：

```
netstat [-a] [-e] [-n] [-s] [-p protocol] [-r] [interval]
```

命令中各参数意义如下。

-a：显示所有连接和侦听端口。

-e：显示以太网统计，可以与-s 选项结合使用。

-n：以数字格式显示地址和端口号。

-s：显示 protocol 指定的协议的统计，默认协议为 TCP、UDP、ICMP 和 IP。

-p：显示 protocol 指定的连接。

-r：显示路由表内容。

interval：重新显示所选的统计，每次显示之间的暂停时间由 interval 设定。

（2）在 Windows 2000 下使用命令行工具 fport。

（3）使用与 fport 功能相同的图形界面工具 Active Ports。

### 3. 常见木马的手工清除方法

不同的木马根据其工作原理和危害有不同的手工清除方法，这些方法在网络上可以搜索到。

### 4. 实验准备

（1）上网搜索一种可以查看开放端口的工具，记录安装和使用方法。

（2）上网搜索并记录多种木马的手工清除方法。

#### 5. 实验内容

（1）下载并安装一种查看开放端口的软件。
（2）用下载的软件查看自己的计算机的端口。
（3）记录查看软件的显示内容。
（4）对所发现的木马进行手工清除。

#### 6. 推荐的分析讨论内容

（1）你遇到过哪些木马？它们有哪些危害？
（2）你使用过哪几种端口查看命令或软件？它们各有什么特点？
（3）其他发现或想到的问题。

### 1.1.4 陷门与黑客攻击

#### 1. 陷门及其特征

陷门（trap doors）也称为后门，是一种操作系统的无口令入口，由一段程序实现，通常是系统开发者为调试、测试、维修而设置的简便入口，操作系统提供的调试器（debug）、向导（wizard）以及 daemon 软件，都有可能被攻击者利用进入系统；有些则是入侵者在完成入侵目标后，为了能够继续保持对系统的访问特权挖空心思设计的，形成隐蔽的信道监视系统运行或伺机对系统发起攻击。所以陷门可以看作人为漏洞。

陷门一般有如下一些技术或功能特征。

（1）陷门通常寄生于某些程序（有宿主），但无自我复制功能。

（2）陷门可以在系统管理员采取了增强系统安全措施（如改变所有密码）的情况下，照样进入系统。

（3）陷门可以使攻击者以最短的时间再次进入系统，而不是重新挖掘漏洞。

（4）许多陷门可以绕过注册直接进入系统，或者帮助攻击者隐藏其在系统中的一举一动。

（5）陷门可以把再次入侵被发现的可能性降至最低。

#### 2. 陷门分类

1）账户与注册陷门

（1）Login 陷门。在 UNIX 中，Login 程序常用来对通过远程登录来的用户进行口令验证。入侵者获取 Login 的源代码并修改，使它在比较输入口令与存储口令时先检查陷门口令。当入侵者输入陷门口令后，Login 程序将忽视管理员设置的口令而让入侵者长驱直入。使用这种方法，入侵者可以进入任何账号，甚至是 root 目录。

（2）密码破解陷门。这是入侵者使用的最老的方法。通常，入侵者寻找口令薄弱的未被使用的账号进行破解，之后将口令改得难一些，形成一些新账号。这些新账号将成为重新侵入的后门。当管理员寻找口令薄弱的账号时，也不会发现这些密码已修改的账号。因而管理员很难确定查封哪个账号。

（3）超级账户陷门。超级账户（Administrator，Admin）是系统安全的宝贵资源。入侵

者一旦可以创建这样的账号，就可以拥有很大的权力。在 Windows NT/2000 上可以使用下面的命令创建本地特权账号：

```
net user <username> <password>/ADD
net localgroup <groupname> <username>/ADD
```

在 UNIX 下，在口令文件中增加一个 UID 为 0 的账户，是创建超级账户的最简单方法。

（4）rhosts++陷门。在联网的 UNIX 机器中，像 Rsh 和 Rlogin 这样的服务是基于 rhosts 的。入侵者只要向可以访问的用户的 rhosts 文件中输入++，就可以允许任何人从任何地方进入这个账户。这些账户也成了入侵者再次侵入的陷门。

2）通信与连接陷门

（1）网络通信陷门。入侵者不仅想隐匿在系统里的痕迹，而且也要隐匿他们的网络通信后门。这些网络通信后门有时允许入侵者通过防火墙进行访问。有许多网络后门程序允许入侵者建立某个端口号，并且不通过普通服务就能实现访问。因为这是通过非标准网络端口的通信，管理员可能忽视入侵者的足迹。这种后门通常使用 ICP、UDP 和 ICMP，但也可能是其他类型的报文。

（2）TCP Shell 陷门。TCP Shell 陷门建立在防火墙没有阻塞的高位 TCP 端口，例如，可能建立在 SMTP 端口，因为很多防火墙允许 E-mail 通过。TCP Shell 后门可以让入侵者躲过 TCP Wrapper 技术。这些端口在许多情况下受口令保护，以防管理员连接后察觉。

（3）UDP Shell 陷门。UDP 是无连接的，管理员不会像观察 TCP 连接的怪异情况那样发现它，因而不能用 netstat 显示入侵者的访问痕迹。所以入侵者通常将 UDP Shell 后门放置在 DNS 端口，因为许多防火墙设置成允许类似 DNS 的 UDP 报文的通行。

（4）ICMP Shell 陷门。由于许多防火墙允许外部 ping 内部的主机，所以入侵者可以将数据放入 ping 的 ICMP 包，在 ping 的主机间形成一个 Shell 通道。虽然管理员也许会注意到 ping 包，但他不查看包内数据就不会了解哪些是入侵者的数据包。

（5）Telnet 陷门。当用户通过 Telnet 连接到系统后，监听端口的 inetd 服务会接受连接并传递给 in.telnetd，由它运行 login。一些入侵者知道管理员会检查 login 是否被修改，就着手修改 in.telnetd。在 in.telnetd 内部有一些对用户信息的检验，比如用户使用了何种终端。入侵者可以对某些服务做这样的后门，对来自特定源端口的连接产生一个不要任何验证的 Shell。

（6）校验和及时间戳陷门。早期，许多入侵者用自己的木马程序替代二进制文件，系统管理员便依靠时间戳和系统校验和的程序辨别一个二进制文件是否已被改变，如 UNIX 里的 sum 程序。为此，入侵者又开发了使木马文件和原文件时间戳同步的新技术。这种技术是这样实现的：先将系统时钟拨回到原文件时间，然后调整木马文件的时间为系统时间。一旦二进制木马文件与原文件精确同步，就可以把系统时间设回当前时间。例如，sum 程序是基于 CRC 校验的，很容易骗过。另外，入侵者设计出了可以将木马的校验和调整到原文件的校验和的程序。不过，MD5（见第 2.2.3 节）是被大多数人推荐的，MD5 使用的算法目前还没人能骗过。

3）隐匿陷门

（1）隐匿进程陷门。入侵者通常想隐匿他们运行的程序。这样的程序一般是口令破解程序和监听程序（sniffer）。有许多办法可以实现他们的目的，较通用的有以下几种方法。

① 编写程序时修改自己的 argv，使它看起来像其他进程名。

② 将监听程序改名再执行。

③ 修改库函数致使 ps 不能显示所有进程。

④ 将一个后门或程序嵌入中断驱动程序，使它不会在进程表中显现。

（2）文件系统陷门。入侵者常要在服务器上存储他们的掠夺品或数据，不希望被管理员发现。入侵者的文件一般有 exploit 脚本工具、陷门集、sniffer 日志、E-mail 的备份和源代码等。为了防止管理员发现这些文件，入侵者需要修补 ls、du 和 fsck 以隐匿特定的目录和文件。在很低的级别，入侵者制作这样的漏洞： 以专有的格式在硬盘上隔出一部分，且表示为坏的扇区，使管理人员难发现这些“坏扇区”里的文件。

（3）共享库陷门。几乎所有的 UNIX 系统都使用共享库，并且管理员很少检查这些库，因此一些入侵者在 crypt.c 和_crypt.c 等函数里做了陷门。

（4）Cronjob 陷门。UNIX 上的 Cronjob 可以按时间表调度特定程序的运行。例如，入侵者可以加入陷门 Shell 程序，使它在深夜 1～2 点运行。那么每晚有一个小时可以获得访问，也可以查看 Cronjob 中经常运行的合法程序，同时植入后门。

（5）内核陷门。内核陷门可以使库躲过高级校验，甚至连静态连接也大多不能识别。

（6）Boot 块陷门。一些入侵者将一些陷门留在根区，因为在 UNIX 下多数管理员不检查根区的软件。

（7）网络服务陷门。网络服务陷门是以服务方式启动陷门或把陷门放置在服务程序有关的文件中。例如在 Windows NT 中，可以采用以服务方式启动陷门程序，使陷门随系统运行而自动启动。这样会给攻击者带来手工不易删除和隐藏性好的好处。在 UNIX 中，由于系统管理员在一般情况下不经常检查超级服务器守护进程（inetd），因此这些配置文件就成为放置陷门的好地方。

### 3. 一些常见陷门工具

下面是黑客常用的创建陷门的工具。

① rootkit、cron、at、secadmin、Invisible Keystoke、remove.exe、rc（UNIX）。

② Windows 启动文件夹。

③ sub7（http://www.sub7.net）。

④ Netcat（http://www.atstake.com/research/tools）。

⑤ VNC（http://www.alvnc.com）。

⑥ BO2K（http://www.sourceforge.net/projects/bo2k)。

### 4. 黑客及其攻击过程

“黑客”一词是对于网络攻击者的统称。一般说来，黑客是一个精通计算机技术的特殊群体。从攻击的动机看，可以把黑客分为 3 类：一类称为“侠客”（Hacker），他们多是好奇

者和爱出风头者；一类称为“骇客”（Crackers），他们是一些不负责的恶作剧者；一类称为“入侵者”（Intruder），他们是有目的的破坏者。随着 Internet 的普及，黑客的活动日益猖獗，对社会造成巨大损失。

黑客进行网络信息系统攻击的主要工作流程是：收集情报、远程攻击、远程登录、取得权限、留下后门、清除日志，主要内容包括目标分析、文档获取、破解密码、日志清除等。这些内容都包括在黑客攻击的 3 个阶段——准备阶段、实施阶段和善后阶段中。

1）攻击的准备阶段

（1）确定目的。一般说来，入侵者进行攻击的目的主要有 3 种类型：破坏型、获取型和恶作剧型。破坏型攻击指破坏攻击目标，使其不能正常工作，主要的手段是拒绝服务攻击（DoS）。获取型主要是窃取有关信息，或获取不法利益。恶作剧型则是进来遛遛，以显示自己的能耐。目的不同，所采用的手段就不同。

（2）踩点，即寻找目标。

（3）查点。搜索目标上的用户、用户组名、路由表、SNMP 信息、共享资源、服务程序及旗标等信息。

（4）扫描。自动检测计算机网络系统在安全方面存在的可能被黑客利用的脆弱点。

（5）模拟攻击。进行模拟攻击，测试对方反应，找出毁灭入侵证据的方法。

2）攻击的实施阶段

（1）获取权限。获取权限往往是利用漏洞进行的。系统漏洞分为远程漏洞和本地漏洞两种，远程漏洞是指黑客可以在别的主机上直接利用该漏洞进行攻击并获取一定的权限。这种漏洞的威胁性相当大，黑客的攻击一般都是从远程漏洞开始的。但是利用远程漏洞获取的不一定是最高权限，而往往只是一个普通用户的权限，这样常常没有办法做黑客们想要做的事。这时就需要配合本地漏洞来把获得的权限进行扩大，常常是扩大至系统的管理员权限。

（2）权限提升。有时获得了一般用户的权限就足以达到修改主页等目的了，但只有获得了最高的管理员权限之后，才可以做诸如网络监听、清理痕迹之类的事情。要完成权限的扩大，不但可以利用已获得的权限在系统上执行利用本地漏洞的程序，还可以放一些木马之类的欺骗程序来套取管理员密码，这种木马是放在本地套取最高权限用的，而不能进行远程控制。

（3）实施攻击。如对一些敏感数据的篡改、添加、删除和复制，以及对敏感数据的分析，或者使系统无法正常工作。

3）攻击的善后工作

（1）修改日志。如果攻击者完成攻击后就立刻离开系统而不做任何善后工作，那么他的行踪将很快被系统管理员发现，因为所有的网络操作系统一般都提供日志记录功能，会把系统上发生的动作记录下来。为了自身的隐蔽性，黑客一般都会抹掉自己在日志中留下的痕迹。为了能抹掉痕迹，攻击者要知道常见的操作系统的日志结构以及工作方式。

（2）设置后门。一般黑客都会在攻入系统后不止一次地进入该系统。为了下次再进入系统时方便一点，黑客会留下一个后门，特洛伊木马就是后门的最好范例。

（3）进一步隐匿。只修改日志是不够的，因为百密必有一疏，即使自认为修改了所有的日志，仍然会留下一些蛛丝马迹。例如，安装了某些后门程序，运行后也可能被管理员发现。所以，黑客通过替换一些系统程序的方法来进一步隐藏踪迹。这种用来替换正常系统程序的黑客程序称为 rootkit，这类程序在一些黑客网站可以找到，比较常见的有 LinuxRootKit，现在已经发展到了 5.0 版本了。它可以替换系统的 ls、ps、netstat、inetd 等一系列重要的系统程序，当替换了 ls 后，就可以隐藏指定的文件，使得管理员在使用 ls 命令时无法看到这些文件，从而达到隐藏自己的目的。

### 1.1.5 其他恶意代码

除了上述介绍的几种恶意代码此之外，属于恶意代码的还有如下多种。

#### 1. 逻辑炸弹

逻辑炸弹（logic bomb）是嵌入某些合法程序的一段代码，没有自我复制功能，通常被预置于较大的程序中，等待某随机事件发生触发其破坏行为。

#### 2. 细菌

细菌（germ）是一种在计算机系统中不断进行自我复制的程序，它们通过不断复制来占有系统资源。细菌也具有独立性。这两点与蠕虫相同，但是，蠕虫一般要利用一些网络工具进行繁殖，而细菌可以自己繁殖。

#### 3. 恶意广告

恶意广告是采用强制弹出、欺骗安装等形式的网络广告，典型的有“插屏流氓”“千尺游戏大厅”推荐的应用和恶意积分墙等。

#### 4. Cookie 与网络臭虫

Cookie（小甜饼）是一种由服务器生成、发送到用户端（一般是浏览器）的文本文件，专门用于保存登录过服务器网站的用户名、密码、浏览过的网页、停留的时间等信息，以便提高网站和用户之间的交互效率。这样，当该用户再次访问同一网站时，服务器就可以决定不要用户输入用户名和密码直接进入已经登录状态，甚至还会根据用户的访问历史主动提供有关信息。Cookie 在生成时就会被指定一个 Expire 值，这就是 Cookie 的生存周期，在这个周期内 Cookie 有效，超出周期 Cookie 就会被清除。有些页面将 Cookie 的生存周期设置为 0 或负值，这样在关闭浏览器时就马上清除 Cookie，不会记录用户信息，更加安全。通常，Cookie 是没有危害的。但是有些互联网企业为了发现“商机”，把“小甜饼”变成了“网络臭虫”（Web bug，也称为网络信标——Web beacon）。“网络臭虫”是一段恶意代码，它们用于收集用户信息、用户访问过的网页、停留时间、购买的商品等个人偏好信息，通过统计分析这些个人信息，向用户精准投放广告，或者再向其他需要这些个人信息的公司出售获利。

### 5. 各种黑客专用工具

一般说来，众多的恶意代码都是黑客的帮凶，或者是黑客利用的工具。不过有一些恶意代码是黑客利用最多的工具。陷门就是其一。此外还有一些是黑客刻意研发出来的工具，例如下列一些。

（1）僵尸网络。僵尸网络与后门类似，也允许攻击者访问系统。但是所有被同一个僵尸网络感染的计算机将会从一台控制命令服务器接收到相同的命令。有关作用将在 1.6.3 节介绍。

（2）下载器。这是一类只是用来下载其他恶意代码的恶意代码。下载器通常是在攻击者获得系统的访问时首先进行安装的。下载器程序会下载和安装其他的恶意代码。

（3）间谍软件。这是一类从受害计算机上收集信息并发送给攻击者的恶意代码。例如嗅探器、密码哈希采集器、键盘记录器等。这类恶意代码通常用来获取 E-mail、在线网银等账号的访问信息。

（4）启动器。用来启动其他恶意程序的恶意代码。通常情况下，启动器使用一些非传统的技术，来启动其他恶意程序，以确保其隐蔽性，或者以更高权限访问系统。

（5）内核套件。设计用来隐藏其他恶意代码的恶意代码。内核套件通常是与其他恶意代码（如后门）组合成工具套装，来允许为攻击者提供远程访问，并且使代码很难被受害者发现。

（6）勒索软件。设计成吓唬受感染的用户来勒索他们购买某些东西的恶意代码。这类软件通常有一个用户界面，使得它看起来像是一个杀毒软件或其他安全程序。它会通知用户系统中存在恶意代码，而唯一除掉它们的方法只有购买他们的“软件”，而事实上，他们所卖软件的全部功能只不过是将勒索软件进行移除而已。

（7）发送垃圾邮件的恶意代码。这类恶意代码在感染用户计算机之后，便会使用系统与网络资源来发送大量的垃圾邮件。这类恶意代码通过为攻击者出售垃圾邮件发送服务而获得收益。

## 1.1.6 信息系统安全卫士

### 1. 信息系统安全卫士概述

“道高一尺，魔高一丈”。随着计算机技术和网络技术的飞速发展与广泛应用，恶意代码和黑客技术也在不断花样翻新、不断升级。这使得计算机用户不得不安装大量防范软件，并搞得手忙脚乱，应接不暇。不过，这种局面也促成信息安全业的发展，也使得信息系统安全卫士应运而生。表 1.3 列举了国内开发的信息系统安全卫士产品。

**表 1.3 国内开发的信息系统安全卫士产品举例**

| 名　　称 | 存储空间/MB | 适用系统 | | | | 更新情况 | |
|---|---|---|---|---|---|---|---|
| | | 微软 | 安卓 | 苹果 | MAC | 新版本号 | 发布时间 |
| 百度安全卫士 | 10.5 | √ | √ | √ | | 8.2.0.7227 | 2016-09-05 |
| 360 安全卫士 | 60.2 | √ | √ | √ | √ | v11.4 | 2017-03-06 |
| QQ 电脑管家 | 52.5 | √ | √ | √ | √ | v12.3 | 2017-03-19 |

续表

| 名　　称 | 存储空间/MB | 适用系统 | | | | 更新情况 | |
|---|---|---|---|---|---|---|---|
| | | 微软 | 安卓 | 苹果 | MAC | 新版本号 | 发布时间 |
| 金山卫士 | 29.0 | √ | √ | | | 4.7.0.4215 官方版 | 2016-03-15 |
| 火绒互联网安全软件 | 9.7 | √ | | | | v4.0.15.2 官方版 | 2017-03-16 |
| 360 Total Security 中文 | 41.8 | √ | | | | 8.8.0.1077 中文版 | 2016-09-22 |
| 2345 安全卫士 | 10.5 | √ | | | | v2.9 官方版 | 2017-01-08 |
| QQ 电脑管家 Win 10 专版 | 50.9 | √ | | | | 11.8.17899.205 | 2016-08-08 |
| 360 安全卫士 Win 10 专版 | 60.2 | √ | | | | v11.4 官方版 | 2017-03-06 |
| 瑞星全功能安全软件 2016 | 44.7 | √ | | | | v23.01.77.85 官方版 | 2016-12-02 |
| 瑞星安全云终端 | 20.6 | √ | | | | v3.0.0.71 官方版 | 2017-03-02 |
| 瑞星安全助手 | 9.7 | √ | | | | 01.00.02.78 | 2016-08-19 |
| 火绒剑独立版 | 2.0 | √ | | | | 0.1.0.36 | 2016-09-19 |
| 阿里钱盾电脑版 | 19.8 | √ | √ | √ | | v1.0.5.92 官方版 | 2016-12-29 |
| 网站安全狗 | 36.8 | √ | | | | v4.0.16786 官方版 | 2017-01-18 |
| 反黑安全卫士 | 34.5 | √ | | | | v1.0.185 官方版 | 2016-08-22 |
| QQ 安全管家 | 32.1 | √ | | | | v6.8 及 v6.9 | 2012-04-19 |
| 360 文档卫士 | 3.3 | √ | | | | v1.0.0.1001 官方版 | 2017-03-17 |
| 网站安全狗 Apache 版 | 25.9 | √ | | | | v4.0.16776 官方版 | 2017-01-18 |
| 悬镜管家 | 7.0 | √ | | | | v3.2.1.3482 官方版 | 2017-01 |
| 铠甲安全卫士 | 35.7 | √ | | | | v2.0.16.1111 官方版 | 2016-11-27 |
| 联想杀毒软件 | 25.8 | √ | | | | v1.0.8 官方版 | 2016-10-19 |

安全卫士其实就是一种信息安全软件包，它集成了多种安全保护软件，并不断升级，为用户提供一个界面、多种功能的信息安全保护。图 1.10～图 1.12 是其中 3 款的用户界面。




图 1.10　百度安全卫士用户界面

功能。




图 1.11　360 安全卫士用户界面




图 1.12　2345 安全卫士用户界面

## 2. 信息系统安全卫士的功能

目前国内开发出的安全卫士产品很多，各有特色，但一般来说都有如下功能或类似功能。

（1）电脑体检：对计算机进行详细的检查。

（2）木马查杀。

（3）漏洞修复：为系统修复高危漏洞和功能性更新。

（4）系统修复：修复常见的上网设置和系统设置。
（5）电脑清理：清理插件、清理垃圾和清理痕迹并清理注册表。
（6）优化加速：加快开机速度。
（7）电脑救援：自助方案、专家人工在线解决计算机故障问题。
（8）软件管家：安全下载软件和小工具。
（9）功能大全：提供其他各式各样的功能，如开机时间、流量显示、性能体检等。

#### 3. 信息系统安全卫士的选用

目前绝大多数安全卫士都是免费提供的。选择一个合适的安全卫士产品，一般应当从可见与不可见两个方面进行比较。

1）可见方面

（1）功能。目前各种安全卫士的功能略有差异。
（2）适用系统。如表 1.3 所示，有的适用于微软（Windows），有的适用于 Android，还有的适用于苹果手机，有的适用于 MAC 系统。有的有多种适用。
（3）占用的存储空间。
（4）版本情况，是否比较新的版本。
（5）界面情况，是否使用起来方便。

2）不可见方面

不可见方面主要要考虑：开发（运行）该软件的商家的实力强弱。一般说来，实力强的公司，样本库丰富、更新快，服务准确、及时。

## 1.2 窃听型攻击

窃听是指使用专用技术装备秘密窃取侦察目标的声波、信号和图像等信息，从中获得情报的一种手段，是窃听、窃视、窃录、窃照的总称。窃听是伴随着竞争出现的。在人类社会中，军事、外交、商业和政治斗争是竞争最为激烈的领域，窃听行为也主要发生在这些领域。

窃听渠道多种多样，技术形形色色。本节介绍基于通信的 5 类监听（声波监听、电磁波监听、光缆监听、手机监听和共享网络中监听）。

### 1.2.1 世界著名监听案例

#### 1. 樗里疾挖地道监听



图 1.13 樗里疾

据称，有史料记载的最早的窃听事件发生在战国时期。《韩非子·外储说右上》记载，樗（chū）里疾（见图 1.13）是秦惠王的弟弟，他足智多谋，很受秦王宠爱。后来，秦王手下来了一个叫公孙衍的谋士，获得了秦王赏识。为了保住

自己的地位，樗里疾在秦王宫殿下面挖了条地道，每当秦王单独召见公孙衍，樗里疾就潜入地道窃听他们谈话，并最终靠窃听得到的消息挤走了公孙衍。这种窃听方式是非常笨拙的，要受许多条件的限制。

### 2. 美国驻苏大使馆中的美国国徽

1960 年的联合国大会期间，当与会代表争论美国 U-2 间谍飞机在苏联领土上被击落的事件时，美国大使亨利 • 卡波特 • 洛奇决定放手一搏。他拿出一个木制的美国国徽，那是苏美友好协会赠送给美国驻莫斯科大使馆的礼物。他用一个镊子从鹰嘴处取下一个微小的麦克风，从而反击了苏联对美国间谍飞机的谴责。

1945 年 2 月雅尔塔会议期间，4 名苏联少先队员抬着一枚由各种名贵木料拼装而成的美国国徽送给卡里曼。这枚内藏“金唇”（它不需要电池和外来电流，就可以接收到 300m 以内大耗电量振荡器所发出的微波脉冲）的美国国徽（见图 1.14）被后来悬挂在卡里曼办公室的那一刻起，克格勃窃听美国大使的“自白”行动便开始启动，持续了 8 年，送走了 4 任美国大使。这个事件被称为苏联的“金唇行动”。

### 3. 美国的“常春藤之铃”

在冷战最激烈的 1971 年，美国“大比目鱼”号潜艇（见图 1.15）奉命前往鄂霍次克海，执行代号为“常春藤之铃”的行动。潜水员们沿着鄂霍次克海北部水下 120m 的深处艰难地发现一串“禁止靠近”的标记后，找到了一条苏联军事通信电缆。这是一条苏联位于符拉迪沃斯托克的太平洋舰队司令部与彼得罗巴甫洛夫斯克潜艇基地进行联络的电缆。电缆中的信号虽是编码的，但却并未加密。潜水员们在这条电缆上安装了能录下所有谈话的窃听器。窃听器是一个长约 5m、直径约 1.2m 的钢柱，内置包含提供能源的钚电池和录音设备。该录音设备能够在电缆有信号时自动开机，最多可录下长达 150h 的通话。美军潜水员每月下水一次，取回录好的录音带，换上新录音带。五角大楼确信这个计划很成功，随后在其他几处苏联通信电缆上也安装了类似的窃听设备。



图 1.14　美国揭露苏联的“金唇行动”



图 1.15　“大比目鱼”号潜艇

“常春藤之铃”行动延续了 10 年之久。直到 1981 年，美国国安局的一位雇员将这一秘密卖给了莫斯科，该行动才被苏联知晓。莫斯科为了掩护美国线人，声称是在修理被渔船毁坏的电缆过程中误打误撞发现了此事。

**4. 美国驻罗马尼亚大使的“皮鞋窃听案”**

1969 年春，美国驻罗马尼亚大使被监听，保安军官最后在大使左脚皮鞋的鞋后跟里发现了一只大功率苏制 K9R 窃听器，这只皮鞋曾由大使馆里的一个“女佣”拿去修理过。在“修理”过程中，鞋后跟被人剜开，装进了这个质量不到 5.7g 的窃听器（见图 1.16），鞋跟上还挖了一个小孔，使窃听器的麦克风头露出来，在另一个小洞里插着一根钢针。这样，只要“女佣”在夜里把针拔出来就关闭了窃听器，而早上在大使起床前把钢针一踩进去，就又开启了窃听器。

**5. 牙齿中的发报机**

米里亚姆（见图 1.17）在意大利是一位家喻户晓的人物，她拥有全意大利最大的侦探社之一，据说年营业额逾一亿美元，在全球更有超过 1400 名合作伙伴。发生在这位风云人物身上的有趣故事是：她在牙医协助下把窃听器装在丈夫的牙齿里，把其曾夸下海口说她绝对抓不到他“偷腥”把柄的老公变成了前夫。

**6. 尼克松的“水门事件”**

1972 年 6 月 17 日凌晨，有 5 个人企图潜入位于华盛顿水门大厦的美国民主党总部时被警方抓获。他们潜入水门大厦是为了更换此前被安放在民主党总部但已失效的窃听设备，不料当场被擒。经过两年的调查，“水门事件”终于暴露了尼克松在 1972 年利用非法手段赢得总统连任的政治丑闻，尼克松的两名竞选顾问和其他近 30 名政府官员先后遭到起诉，尼克松也在强大的压力下于 1974 年 8 月 9 日辞职，成为美国历史上唯一一位辞职的总统。图 1.18 为尼克松发表辞职演说的电视画面。



图 1.16　皮鞋中的窃听器



图 1.17　意大利女侦探米里亚姆



图 1.18　尼克松发表辞职演说的电视画面

**7. 密特朗窃听门**

弗朗索瓦 • 密特朗（Francois Mitterrand，见图 1.19）于 1981—1995 年任法国总统。1982 年 8 月 18 日，当时的法国总统密特朗下令组建一个“反恐怖合作、信息和行动委员会”，并将这一任务交给了当时的国家宪兵快速行动小组负责人普鲁托。在普鲁托的领导下，法国秘密建立了“爱丽舍宫电话监听系统”。这一系统的初衷是预防恐怖事件的发生，但这一职能很快得到了“延伸”。1982 年 8 月 28 日，普鲁托将窃听系统的触角伸向了某政要的私

人住宅，而这一行动受到了密特朗总统的“理解与支持”。

1997 年 2 月，调查人员在一个汽车修理厂内发现一批普鲁托的个人资料，其中有部分材料显示，前总统密特朗曾发布过有关窃听的命令。经过数年调查确认，该系统 1983—1986 年对 250 位公众人物进行了非法的电话窃听。为此，有 12 名密特朗的前助手成为被告，而密特朗于 1996 年 1 月去世，躲过了司法追究。

### 8. 苏联解体的导火索

1991 年 7 月底，苏联领导人戈尔巴乔夫同当时的俄罗斯总统叶利钦和哈萨克斯坦总统纳扎尔巴耶夫的机密电话被窃听。当时 3 人谈到克格勃首脑克留奇科夫和国防部长亚佐夫的去留，戈尔巴乔夫认为“那些达到退休年龄和构成负担的人应该被替换”。谈话被克格勃秘密偷听并录音，克留奇科夫利用录音胁迫亚佐夫共同发难。1991 年 8 月 19 日，以克留奇科夫为首的“紧急状态委员会”发动政变，政变终因人民反对、军队倒戈而流产，苏联也在政变中宣告解体。图 1.20 为苏联解体过程中的叶利钦指着戈尔巴乔夫的照片。



图 1.19　弗朗索瓦 • 密特朗



图 1.20　叶利钦指着戈尔巴乔夫

### 9. 联合国遭窃听事件

2004 年 2 月 26 日，英国前国际开发事务大臣克莱尔 • 肖特对媒体披露称，在伊拉克战争爆发前，英国负责海外情报事务的军情六处曾经在政府的授权下对联合国秘书长安南的办公室进行窃听，她本人就翻阅过安南私人讲话的监听记录文件副本。

第二天，联合国前首席武器核查官理查德 • 巴特勒也站了出来，从侧面证实了肖特的指控。巴特勒曾经在 1997—1999 年担任联合国负责监督伊拉克销毁大规模杀伤性武器的特别委员会主席。巴特勒对美联社记者说：对他进行间谍活动的其实不止英国一家。他相信在联合国安理会的 5 个常任理事国中，除了中国以外，另外 4 个国家都对他进行了窃听活动。

曾经在 1992—1996 年担任联合国秘书长的加利也向外界表示，他确信在任职期间一直是间谍活动的目标之一。加利在接受英国广播公司的采访时说：“从我上任的第一天起，就有人提醒我，让我小心，因为我的办公室和住宅都被人安装了窃听器。”加利还透露，据他所知，“具备技术手段”的国家有不少都对联合国秘书长进行窃听活动，这已经成了一个“传统”。而联合国对此也几乎无能为力，所以只能自己采取防范措施，因为不光家里、办公室

里，甚至汽车以及电话上都有别人的“耳朵”。一位曾在加利身边工作多年的联合国工作人员称，为了防止遭到窃听，这位前联合国秘书长从上班第一天起便几乎养成了一个习惯：每天上班之前先把自己的衣服抖搂几下，再让手下把文件包、办公室角落悉数“扫描”一遍，再认真地检查一下电话机，看看下面是否有窃听器。

2004 年 12 月，联合国驻欧洲办事处发言人玛丽称，该办事处所在的日内瓦万国宫法兰西厅内发现了一套窃听装置。窃听装置是当年秋天工人们在法兰西厅的装修施工过程中在护墙板内发现的。

2013 年 8 月 25 日，德国《明镜》周刊获得的美国国家安全局秘密文件显示，美国国家安全局不仅监听欧盟，而且也对联合国总部实施了监听行动。

**10.《世界新闻报》窃听丑闻**

2005 年 5 月，《世界新闻报》刊登了威廉王子膝盖肌腱受伤的消息。这条不痛不痒的报道引起了英国王室的怀疑，因为王子受伤的事情鲜有人知，王室随即向警方报案。2006 年 4 月，《太阳报》（*The Sun*）刊登了哈里王子流连脱衣舞夜总会的新闻，随后《世界新闻报》跟踪报道，在刊登的新闻中竟然还原了威廉王子嘲笑哈里的邮件。

英国《太阳报》是富商基思 • 鲁珀特 • 默多克（Keith Rupert Murdoch）拥有的新闻集团（News Corporation）旗下的一份小报。《太阳报》是全英国销量最高的报纸，2004 年后期该报的每日发行量达 320 万份。星期日版的《太阳报》名为《世界新闻报》（*News of the World*）。该报纸以报道许多名人以及涉及名人的丑闻的文章而出名。

伦敦警方于 2006 年开始调查这家报纸窃听名人电话事件。据透露，《世界新闻报》窃听行为的受害者可能多达 4000 人。其中涉及它曾雇私家侦探窃听 2002 年失踪女孩米莉 •道勒手机语音留言。这一系列的监听丑闻，引起从首相到普通民众的一致愤慨。

2011 年 7 月 10 日，《世界新闻报》被迫关门停刊，当日出版的最后一期向读者告别（见图 1.21）。

**11. 美国的“棱镜计划”**

棱镜计划（PRISM）是一项由美国国家安全局（NSA，见图 1.22）自 2007 年小布什时期起开始实施的绝密电子监听计划，该计划的正式名号为 US-984XN。美国情报机构一直在 9 家美国互联网公司中进行数据挖掘工作，从音频、视频、图片、邮件、文档以及连接信息中分析个人的联系方式与行动。监控的类型有 10 类信息：电邮、即时消息、视频、照片、存储数据、语音聊天、文件传输、视频会议、登录时间和社交网络资料的细节，其中包括两个秘密监视项目，一是监视、监听民众电话的通话记录，二是监视民众的网络活动。2013 年 7 月 1 日晚，维基解密网站披露，美国“棱镜门”事件泄密者斯诺登（Edward Snowden，见图 1.23）在向厄瓜多尔和冰岛申请庇护后，又向 19 个国家寻求政治庇护。从欧洲到拉美，从传统盟友到合作伙伴，从国家元首通话到日常会议记录，几乎所有人的通信活动都处于“棱镜计划”监听之中。




图 1.21　最后一期《世界新闻报》



图 1.22　美国国家安全局



图 1.23　爱德华·斯诺登

## 1.2.2　声波窃听

声波是一种波动，因此它具有波动的一切特性，能产生反射、折射、干涉、衍射、共鸣等现象。根据声波特性，人们制造了多种多样的窃听器。

早在2500年前的战国时代，就出现了一种称为“听瓮”的监听工具。听瓮是用陶制成的，如图1.24（a）所示，它大肚小口。把它埋在地下，并在瓮口蒙上一层薄薄的皮革，人伏在上面就可以倾听到城外方圆数十里的动静。到了唐代，又出现了一种“地听器”，如图1.24（b）所示，它是用精瓷烧制而成的，形状犹如一个空心的葫芦形枕头，人睡卧休息时，侧头贴耳枕在上面，就能清晰地听到30里外的马蹄声。



(a) 听瓮



(b) 地听器

图 1.24　听瓮和地听器

北宋大科学家沈括在他著名的《梦溪笔谈》一书中介绍了一种用牛皮做的“箭囊听枕”。他还科学地指出，这种“箭囊听枕”之所以能够听到“数里内外的人马声”，是因为“虚能纳声”，并利用共振来放大传来的微弱信号，而大地又好像是一根“专线”，连接着彼此两个地点，是一种传递声音信号的媒介。在江南一带，还有一种常用的“竹管窃听器”。它是用一根根凿穿内节的毛竹连接在一起的，敷设在地下、水下或隐蔽在地上、建筑物内，进行较短距离的窃听。

在国外，人们还采用过称为“大耳朵”的窃听器（见图1.25）。这种窃听器有一个特别大的圆盘，圆盘朝前的一面为抛物面，当正前方传来的声波碰到圆盘时，根据波的反射原理，会被圆盘反射聚集在焦点上，来自其他方向的声波则不会聚焦。在焦点上放置一个能接收微弱声音的微音器，从正面传来的微弱声音激励微音器工作，将声能转换成电信号，

经电子线路放大，再由窃听人员使用耳机监听。这种抛物面式窃听器能够拾取较大面积的声能，窃听距离可达几千米。

根据同样的声波反射、折射原理，还可制成外形像扩音喇叭一样的远距离定向麦克风窃听器。为了提高灵敏度和指向性，还可根据双耳效应，用两个喇叭拾音。所谓双耳效应，就是来自正前方的声音同时到达双耳，而来自侧面的声音，由于传播路程略有差异，总是一个耳朵先听到，另一个耳朵后听到，分析两耳听到声音的时间差，就可确定声音的方向和距离。

为了便于携带，人们还根据波的叠加原理制成了外形像鸟枪的窃听器，窃听者只要把“鸟枪”的枪口对准被窃听的方向，就能取得较好的窃听效果。这种“鸟枪”窃听器在它长长的枪管上开有很多规则排列的小孔，当声波从正前方传来时，经过小孔进入枪管，就会在枪管尾部的微音器处互相加强；而当无关的声波从枪管两侧传来时，经小孔进入枪管后则互相抵消，这就使监听人员听不到与窃听对象无关的声音，只拾取被侦察方向的声音。

声波是疏密波，在稀疏区域实际压强小于原来的静压强，在稠密区域实际压强大于原来的静压强，声压的周期性变化可以控制电流的周期性变化，从而把声信号转换为电信号，然后经输送线传到电声装置，再将电信号转换为声信号，以供监听人员接收。这种利用声振动产生声压传递信号的原理，不仅是窃听器的工作原理，也是电话机的工作原理。随着电话的普及，电话机也成为常用的窃听工具。最初使用电话进行窃听的方法如图 1.26 所示，是将微小窃听器放到电话的麦克风中，进行声波窃听。



图 1.25 “大耳朵”



图 1.26 放进电话机中的窃听器

为了便于隐藏，窃听器日趋小型化、微型化。有的窃听器做成黄豆粒或针尖那么小，埋设在墙壁、电话机、电灯、沙发、椅子里，用一对导线将信号引出来。窃听者在远远的地方即可听到室内的动静，其拾音范围可达 10m 左右，甚至连写字的声音都能听得一清二楚。为了减少专设线路和解决窃听器的用电问题，往往就利用室内电源插座上的交流电，窃听者只要在电源插座上附设小小的配件，窃听麦克风拾取的谈话声音，经过放大调频变成载波信号，送到电源线上传输出去。窃听者在电源线路的任何位置接上一个载波接收器，便能听到室内的谈话。

为了隐蔽，许多窃听器被隐藏在日常用品中。例如，国外曾有一种伪装成航空卡片架的窃听器，上面用法文写着“航空运输联盟”等字样，但实际上把微型麦克风窃听发射机、遥控接收机和电池等装在不到 1cm 厚的木板底座里，上面的金属框架就是发射天线和接收天线。有的微波窃听器可以制得很小，隐藏在提包、首饰、钢笔、眼镜、鲜花、领带、纽

扣、餐厅服务员临时送上的调料、烟灰缸以及电子打字机、计算机、译码电信机、电传机、保密机等电子设备中。1983 年 1 月，法国驻莫斯科使馆在修理电传打字机的过程中，意外发现打字机的电容器里有个复杂的电子窃听器，它可使电文在未被译成密码之前就被截收。

还有一种方法是在建筑物中预先埋设窃听器。例如，苏美情报机构常常利用为外国修建或改建大使馆的机会，把这种窃听装置埋设在使馆内。20 世纪 80 年代美国曾派特工渗入负责兴建苏联驻美使馆的建筑商，趁机在大使馆地底下挖掘了一条秘密隧道，在隧道中安装各种窃听工具，监视大使馆的一举一动。而当时的苏联也采取了同样手段。1985 年美国在对其驻苏大使馆的新馆舍进行安全检查时，在混凝土构件中查出了一大堆麦克风。1987 年里根说："美国除了全部拆除驻苏新使馆大楼外，别无选择。整幢大楼窃听器密布。"

为了解决入室布放困难的问题，人们也绞尽脑汁。例如，可将拾音器、信号放大电子线路、电池、天线等装在特制的炮弹中，在作战之前伴随火力侦察，发射到敌方的哨所、驻地、指挥部附近，或交通要道等处，它可以起着侦察兵起不到的作用。

克格勃在 20 世纪 50 年代中期广泛应用的一种微型无线电窃听器"虫威"是这个时代窃听器的代表作，其体积只有火柴盒大小，可以用气枪弹射到窃听目标外墙上，用超短波将所收到的声音发射到直径为 5 英里的范围之内，用一个灵敏度很高的接收机就能收到。

美国中央情报局在 20 世纪 60 年代后期采用集成电路生产了一种直径 0.25cm 的无线电传送器，并把它安装在苍蝇背上。执行任务前，他们会让苍蝇先吸入一口神经毒气，使它到达目的地完成窃听器的布放后很快死去。

20 世纪 70 年代初，美国中央情报局利用训练过的鸽子布放窃听器。他们把微型窃听器系在鸽子身上，然后将红色激光束射向要进行窃听的窗户上，鸽子就乖乖按照激光导向，飞落在这个窗户的窗台上。它啄一下按钮，窃听器便脱离鸽身，开始自动工作。

在冷战时期，美国曾进行过一项代号为"声响小猫"的试验。他们在猫体内放入窃听器、电池和线路，猫的尾巴被用作天线。按照中情局原先的设想，他们最终要把这只猫变成可遥控指挥、听话的"高级间谍"。只是由于试验失败，这一计划才被迫告吹。

随着科学技术新成就的不断出现，窃听技术越来越升级，其方法和手段越来越多，如激光窃听、辐射窃听等新的窃听技术相继问世。例如，由于激光可以探测到物体表面极微弱的振动，20 世纪 60 年代激光技术问世不久也被窃听技术专家利用，制作出激光窃听器。这样，室内谈话的声音所引起的窗户上玻璃的轻微振动，可以用激光来对准窗玻璃发射，再用一个激光接收器接收由窗户玻璃反射回来的激光，还原成声音。

可以说，在今天，人们都不知道什么地方没有窃听器了。一个耐人寻味的故事是，1955 年已经有了可以放进手表中的窃听器，当时美国在柏林的特工会见一东德同行时使用的就是这种监听器。尽管事先有一个禁止记录谈话的协定，但谁也没有遵守。没想到在谈话中美方的录音机出了毛病，发出响声。而德方人员还以为是自己偷带的录音机出了故障，甚至尴尬地问："先生，是您的录音机出了毛病呢，还是我的？"

### 1.2.3 电磁波窃听

#### 1. 电磁波窃听的种类

电磁波窃听是截获载有信息的电磁波的窃听手段。它有两种形式：信号拦截监听和电磁泄漏监听。

1）信号拦截监听

信号拦截监听也称为搭线窃听，其方法是在被监听的信道加装信号拦截装置。例如，将窃听器的两根接线接到电话线路上，直接截获电话线路里的电流信号。“水门事件”就属于典型的搭线窃听。为了隐蔽，窃听者常把窃听位置选择在电话线路的接线盒内、分线箱上，尽量不入侵室内。现在已有自动化程度很高的旁听设备，一旦有人拿起手机准备打电话，电话集中台便自动开始工作，数字显示器就显示出该电话机的号码，自动报时器报告通话开始和结束的时间，录音机录下电话内容。此外，还可根据电磁感应现象，将感应线圈设置在电话线外、电话机下，以此来窃听电话内容。

2）电磁泄漏监听

电磁泄漏是指电子设备中的杂散能量向外扩展，并在扩展过程中夹带了设备所处理的数据信号。在一定的条件下，在一定的距离内，重新复原这些信息已经不是难事。

1985 年，在法国召开的一次国际计算机安全会议上，年轻的荷兰人范•艾克当着各国代表的面，用价值仅仅几百美元的器件对普通电视机进行改造，在楼下的汽车内，接收并显示出了 8 层楼上计算机屏幕上显示的图像。国外也有实验表明，银行计算机上显示的密码，竟在马路上就轻易地被截获了。

实际上，计算机的数字信号就是一些高频脉冲信号。这些高频电磁信号会产生与一台小型电台差不多的电磁波辐射。

最早用来捕获电磁波泄漏信号的设备是矿石无线电收报窃听器。矿石收音机（见图 1.27）是最简单的无线电接收机，由长导线天线加上选择信号频率的调谐器和检波器组成，因为检波器可以使用晶体矿石，所以称为矿石收音机。据说矿石收音机至今可能依然被间谍使用，因为它没有振荡器，不需要电池和电能，因此反间谍组织不能侦测到被监听的频率。

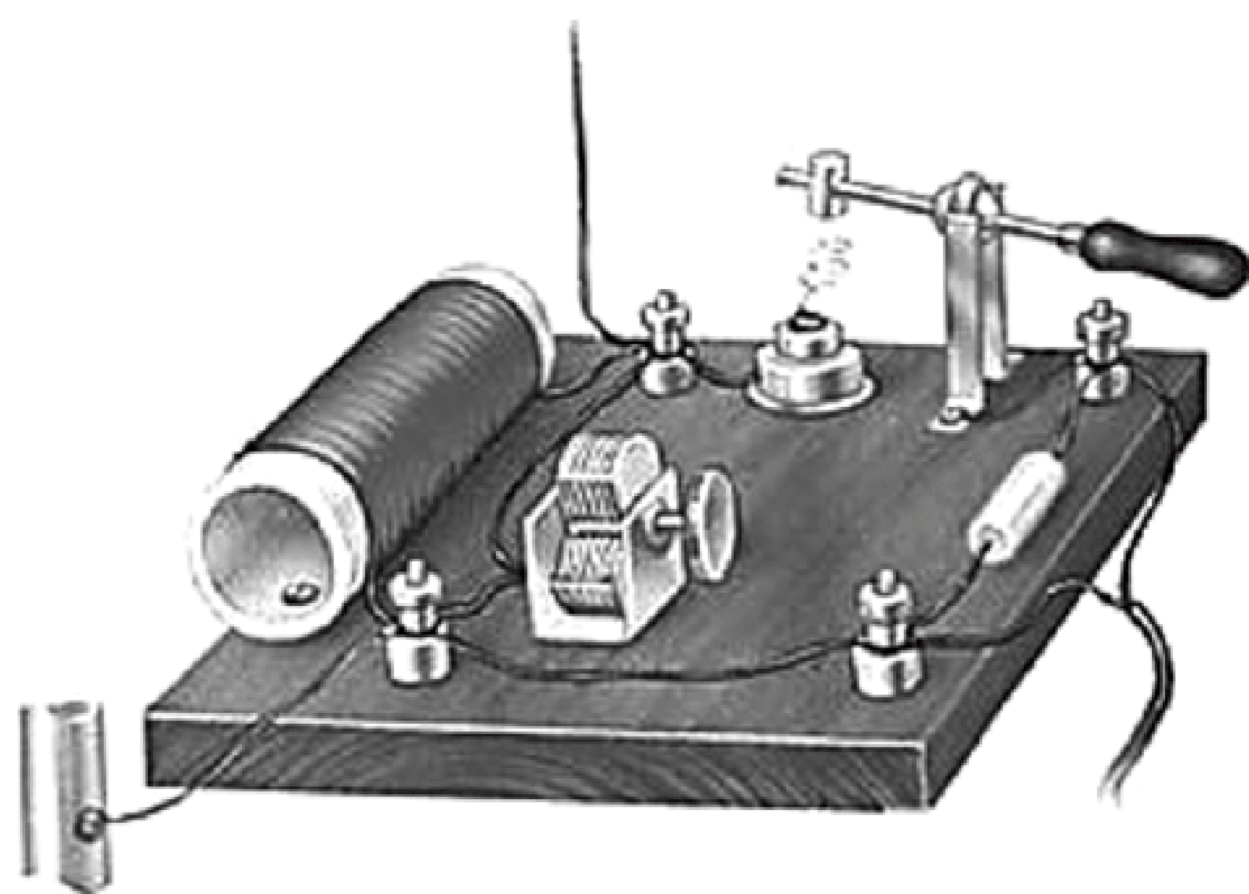


图 1.27　矿石收音机

1914 年夏天，第一次世界大战时期，在法国北部一座幽静的花园里停着一辆毫不起眼的马拉大篷车。这辆车里安装着当时英国军事情报局最先进的矿石无线电收报窃听器，用它来窃听邻近德国军队的无线电联系信号。

随着天线技术的进步，人们已经可以捕捉到距离更远、强度更弱的电磁波信号了。据美军一份泄密文件透露，驻日美军楚边基地是美国家安全局遍布全球的情报网的重要一环，

楚边通信所始建于 1957 年，于 1962 年完工。在冷战时期，美国国家安全局能够监听到苏联的所有密码通信，冷战结束后，由于所处地理位置的独特性，楚边通信所非但没有被拆除，反而在对朝鲜半岛以及中国和东南亚地区的侦察中发挥着不可替代的作用。图 1.28 为位于冲绳美军基地的楚边通信所内，用来搜集周边国家情报的天线“象栏”。图 1.29 为日本在三泽基地密布的球形天线。



图 1.28　美军部署在日本的巨形雷达天线“象栏”



图 1.29　日本在三泽基地密布的球形天线

随着大数据处理技术日臻成熟，广泛进行电磁泄漏信号的监听往往可以获得意想不到的有价值信息。

**2. 电磁波窃听的防范**

针对电磁泄漏可以采取如下一些对抗措施。

（1）屏蔽：用电磁屏蔽技术，既可防止电磁波外泄，又可防止外来电磁波的干扰。

（2）隔离：将需要重点防护的设备从系统中分离出来，加以特别保护。

（3）使用低辐射计算机设备，在分立元器件、集成电路、连线器和阴极射线显示器等方面采取防辐射措施。

（4）使用干扰器：产生电磁噪声，增大辐射信息被截获后破解还原的难度。

（5）滤波：在电源或信号线上加装合适的滤波器，阻断传导泄露的通道。

（6）接地：接地可以使杂散能量向大地泄露。

（7）数据加密和数据隐藏。前者隐蔽了数据的可读性，后者隐蔽了数据的可见性。

### 1.2.4　光缆窃听

一直以来，人们都认为在电缆中传输的信息很容易通过搭接方式窃取，并且由于当电缆有电流通过时，在导体周围会产生磁场，设备足够灵敏就能感应到这个磁场的变化，而无须物理分割导体的金属载体，而光缆传输则非常安全。

然而，2005 年 3 月，美国“海狼”级核潜艇“吉米·卡特”号（见图 1.30）的服役，彻底改变了人们的认识，这是因为“吉米·卡特”号潜艇是一艘专攻海底光缆窃听的潜艇。该潜艇具备海底光缆窃听功能，配备有专门用于安装窃听装置的深潜器，其最大下潜深度达到 610m，通过坐沉海底，释放深潜器实施窃听，或者将窃听装置安装到光缆上长期监听。

实际上，美国早在 20 世纪 90 年代中期就进行过光缆窃听试验，目前的光缆窃听技术已经十分纯熟。据美方资料透露，对光缆进行窃听主要有两种方式：光纤窃听和中继站窃听。

对光缆进行窃听的技术也不止一种。一种技术是将含有一根一根光纤的极细“针管”插入光缆内护套，直抵光纤。于是，针头中的光纤与光缆中的光纤连接，光束会被部分引入窃听装置。而光缆中的光强衰减并不影响光缆正常工作。另一种方法是将光缆剥开至裸纤，将光缆弯曲到临界角度，使得一部分光线从光纤中折射出来。其基本方法如图 1.31 所示，从光纤中折射出来的光线被设备中的光学检测设备拾取，然后发送给光电转换设备将光信号转换为电信号，再将这些信号送计算机分析。




图 1.30　美国“海狼”级核潜艇“吉米·卡特”号

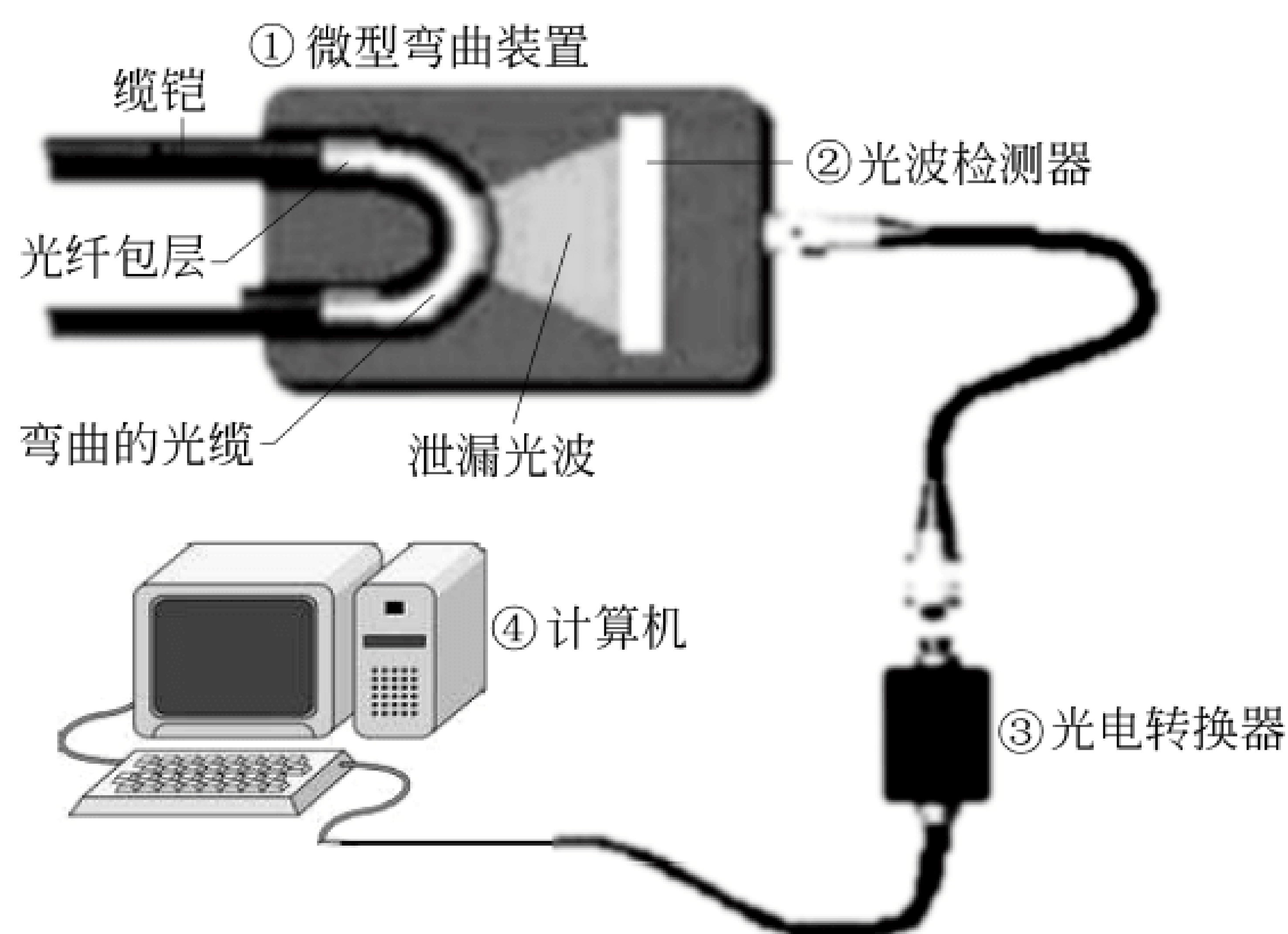



图 1.31　弯曲光缆进行窃听

光纤对比法也是美国较常用的窃听方式，让与激光不同波段的光线沿光纤的径向射过光纤，从而得到相应脉冲信号的光信号，进而转换成电信号，达到窃听目的。

而中继站窃听更为容易，即通过打开光缆中继器加装窃听装置实现窃听。

不过，由于大多数光纤窃听技术都会导致光纤内部光束能量的微小减弱，因此检测光纤能量衰减就是一种窃听发现技术，其中就包括宽波带能量监测，当监测到的通信服务下降达到超过一定阈值时，就认为遭到攻击。

对付光缆窃听，除了检测能量衰减外，采用量子通信是被广泛认可的技术。量子通信是指利用“量子纠缠”效应进行信息传递的一种新型通信方式。所谓“量子纠缠”，是指在微观世界里，不论两个粒子间距离多远，一个粒子的变化都会影响另一个粒子的现象。这个现象被爱因斯坦称为“诡异的互动性”。作个形象的比喻，纠缠状态下的量子就像一对“心有灵犀”的骰子。甲、乙两人身处两地，各拿其中一个骰子，甲随意掷一下骰子是 5 点，与此同时，乙手中的骰子就自动翻转到 5 点。

量子通信涉及量子密码通信、量子远程传态和量子密集编码等。基于量子力学的基本原理，量子通信具有高效和绝对安全等特点，因此成为国际上量子物理和信息科学的研究热点。这门学科已逐步从理论走向实验，向实用化发展。可喜的是，我国对于量子通信技术的研究已经取得了重大进展。据报道，我国正在研制可以自行毁灭的量子密钥，严防他国间谍试图侦察、窃听和窃取量子加密的数据。目前，我国的广域量子通信网络计划已经开始实施，未来 2～5 年，量子通信技术将得到广泛拓展应用。

### 1.2.5　共享网络中的窃听

现在实际运行的计算机网络基本上是一种如图 1.32 所示的“TCP/IP＋ 以太网”结构。在这种网络中，当有两台主机通信的时候，源主机（记为 A）发往目的主机（记为 B）的数据包，要先在运输层（TCP/UDP）标记上端口（进程）号，在网际层加上主机的 IP 地址，成为网际层数据包。但是，这种数据包不能在 IP 层直接传送，必须再交给网络接口，在物理网中传送。然而，物理网不会识别 IP 地址，还必须添加以太帧头信息，分别为只有物理网才能识别的源主机和目的主机的物理地址，它们是与 IP 地址相对应的 48 位的地址——MAC 地址。对于作为网关的主机，由于它连接了多个网络，该 MAC 地址对应着很多个 IP 地址，即该 MAC 地址在每个网络中都有一个对应的 IP 地址，其他主机则是一个物理地址对应一个 IP 地址。

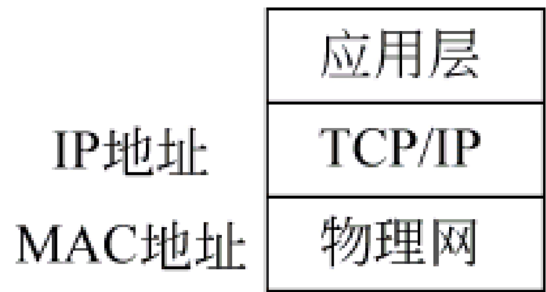


图 1.32　现行网络基本结构

由于现在的物理网多为以太网，所以每台主机的 MAC 地址实际上是其网卡的编号，这个网卡号是全世界唯一的。

网卡具有如下 4 种工作模式。

（1）广播模式（broadcast mode）。它的目的地址是 0Xffffff 的帧为广播帧，工作在广播模式的网卡接收广播帧。

（2）多播模式（multicast mode）。多播传送地址作为目的 MAC 地址的帧可以被组内的其他主机同时接收，而组外主机却接收不到。但是，如果将网卡设置为多播传送模式，它可以接收所有的多播传送帧，而不论它是不是组内成员。

（3）直接模式（direct mode）。工作在直接模式下的网卡只接收目的地址匹配本机 MAC 地址的数帧。

（4）混杂模式（promiscuous mode）。工作在混杂模式下的网卡接收所有流过网卡的帧，数据包捕获程序就是在这种模式下运行的。

网卡的默认工作模式包含广播模式和直接模式，即它只接收广播帧和发给自己的帧。如果采用混杂模式，一个站点的网卡将接收同一网络内所有站点发送的数据包，这样就可以达到对网络信息进行监视和捕获的目的。

早期的以太网采用的是总线结构，即各台主机使用集线器（hub）连接。这时，数据帧可以到达每一台主机。当网卡在默认模式下工作时，如果到达的数据帧中携带的物理地址是自己的或者是广播地址，则将数据帧交给上层协议软件——IP 层软件处理，否则就将这个帧丢弃，即数据帧只能被与目的地址相符的主机接收。但是，若网卡的工作模式被设置成混杂模式，这台主机就可以接收所有到达的数据帧了。在这样的网络中实施监听并非难事，并且监听可以在网上的任何一个位置实施，如局域网中的一台主机、网关上或远程网的调制解调器之间等。

在 UNIX 系统上，当拥有超级权限的用户要想使自己所控制的主机进入监听模式，只需要向 Interface（网络接口）发送 I/O 控制命令，就可以使主机设置成监听模式了。而在 Windows 9x 的系统中，则不论用户是否有权限，都可以通过直接运行监听工具进入监听模式。

### 1.2.6 手机监听

#### 1. 手机监听概述

随着手机成为人们生活中不可或缺的必需品，窃听开始与手机紧密相连，手机监听的事件也层出不穷。1996 年 4 月，俄罗斯车臣叛乱分子的头目杜达耶夫因手机泄密，被俄军发射导弹击毙。2002 年 3 月本•拉登的得力助手、“基地”组织的二号人物阿布•祖巴耶达赫因使用手机暴露藏身地而落网。2013 年 10 月 23 日，德国媒体报道了美国情报部门监听德国总理默克尔的消息，如晴空霹雳，一时间舆论大哗，默克尔无限委屈，恼怒万分（见图 1.33）。



图 1.33　得知手机被美国情报部门监听的默克尔

一般说来，手机监听的技术有如下两种。

1）截获手机的电磁波

一般说来，手机的通信过程就是使用手机把语音信号传输到移动通信网络中，再由移动通信网络将其变成电磁频谱，通过通信卫星辐射漫游传送到受话人的电信网络中；受话人的通信设备接收到无线电磁波，再转换成语音信号接通通信网络。手机使用的无线信道的开放性，让第三者只要有相应的接收设备，就能够截获任何时间、任何地点、任何人的通话信息。拦截距离可达上万千米，有效监听距离与地球同步通信卫星信号覆盖范围几乎相等。

2）安装手机“卧底”软件——木马程序

强大的卧底软件可以完成如下一些监控。

（1）隐秘地进行环境音效拦截监控。如果目标手机是空闲状态，通过拨打使这个手机被秘密接通，卧底软件会自动激活目标手机的免提麦克风而不会有任何显示，目标手机可以清楚地传回周围环境的声音。如果目标手机正在使用中或者目标手机的使用者按了任意键，本次呼叫将会被秘密断开，不留一点痕迹。

（2）对目标手机进行定位追踪。对于有 GPS 装置的手机，卧底软件可以清楚地确定目标手机的经纬度坐标以及定位时完整的地图显示。有些手机虽然没有 GPS 装置，但是所有手机都有电话身份识别功能，手机卧底软件还可以提供基于移动基站查询定位的功能。

（3）短信监控。手机卧底软件可以根据设置，捕获目标手机发送、接收的短信，使窃听者可以在远端查看到短信的内容、对方号码、发送/接收时间等信息，如果对方号码在目标手机的通讯录（联系人）存有姓名，对方号码会与姓名关联，那么还会显示对方的这个名字。

（4）电子邮件监控。手机卧底软件会根据设置，捕获目标手机所发送、接收的电子邮件信息并上传到手机卧底服务器。

（5）远程遥控。手机卧底软件可以使用指令对目标手机进行遥控设置，如更改监控号码、更改上传频率、发送诊断请求等。

（6）手机卧底软件让目标手机永不脱离窃听者的视线。当目标手机更换 SIM 卡（手机卡）后，手机卧底软件会使用目标手机新更换的 SIM 卡以秘密短信的形式把监控号码发送给监控者，告知 SIM 卡已变更的信息。监听者可以根据此信息，对目标手机新的号码进行监控。

有一些监听是根据手机网络的制式特点进行的。下面讨论针对 GSM 和 CDMA 这两种常用手机制式的监听原理。

**2．GSM 手机监听**

1）利用 GSM 信道漏洞监听

全球移动通信系统（Global System for Mobile Communications，GSM）是世界上主要的蜂窝系统之一。GSM 是采用 FDMA（频分）与 TDMA（时分）制式相结合的一种通信技术，其网络中所有用户分时地使用不同频率进行通信。中国 GSM 手机采用 900MHz 频段和 1800MHz 频段工作。表 1.4 为 GSM 工作频段的具体分配。

**表 1.4　中国 GSM 工作频段分配**

| 种　类 | 上行频段/MHz | 下行频段/MHz |
|---|---|---|
| GSM900 | 890～915 | 935～960 |
| GSM1800 | 1710～1785 | 1805～1880 |

对于 GSM900，上下行频段的各 25MHz 的频率范围被划分为 124 个不同的信道，每个信道带宽为 200kbps，每个信道采用 TDMA 技术分为 8 个时隙，形成 8 个物理信道，理论上一个射频允许同时进行 8 组通话。所以 GSM900 频段在同一区域内可同时供近 1000 个用户使用。

GSM900 的帧时长为 4.615ms，每个物理信道的时隙长度为 0.577ms，即把时间分割成周期性的帧，每一帧再分割成许多个时隙。之后根据特定的时隙分配原则，使移动手机用户在每帧中按指定的时隙向基站发送信号。基站分别在各自指定的时隙中，接收到不同的移动手机用户的信号，同时基站也按规定的时隙给不同的移动手机用户发射信号。各移动用户在指定的时隙中接收信号。为保证在同一信道上的用户可以相互不受干扰，GSM 按欧洲和亚洲的应用标准采取 5 级保护。

（1）用户接入网络时的鉴权。

（2）移动设备识别。

（3）无线路径信号加密。

（4）临时识别码保护。

（5）以个人身份识别码（Personal Identification Number，PIN）保护 SIM 卡。

如果电信运营商真的严格执行了这些标准，就会大大增加监听难度。但现实中，有些运营商往往出于利益目的，使这些标准的实施有可能不完全到位，形成一些漏洞。例如，当一个人使用 GSM 进行通信的时候，其手机和 GSM 网络将对本次会话用一个临时 ID 和会话密钥进行加密。但是，由于 GSM 的加密算法存在缺陷并重复地使用相同的会话密钥，这样，如果数据被记录，黑客会很快而且很容易解密会话密钥和临时 ID，然后黑客可以使用

临时 ID 和会话密钥去伪造该号码的通信。

2）利用接口漏洞监听

从信令（signal，电信网中的控制信号）结构上看，GSM 系统包括如下一些接口。

（1）移动应用部分（Mobile Application Part，MAP）接口。

（2）A 接口（GSM 网络子系统 NSS 与基站子系统 BSS 之间的标准接口）。

（3）ABIS 接口（基站控制器 BSC 与基站收发信台 BTS 之间的通信接口）。

（4）UM 接口（GSM 的空中接口——基站与移动台间的接口）。

这些接口都有大量的性能参数和配置参数，一些具体参数在设备完成前就已经设定好了，这里面本身就存在许多漏洞。另外，一般人不知道的是一些国外生产的手机也都是留有监听接口的。

3）利用其他漏洞监听

GSM 还有一个特点是其发射功率较大，这也为远程监听提供了一些条件。

**3．CDMA 手机监听**

码分多址（Code Division Multiple Access，CDMA）是在扩频通信技术上发展起来的一种崭新而成熟的无线通信技术。如图 1.34 所示，它将需传送的具有一定信号带宽的数据用一个带宽远大于信号带宽的高速伪随机码进行调制，使原数据信号的带宽被扩展，再经载波调制并发送出去。接收端使用完全相同的伪随机码对接收的带宽信号做相关处理，把宽带信号换成原数据的窄带信号，即解扩，以实现信息通信。

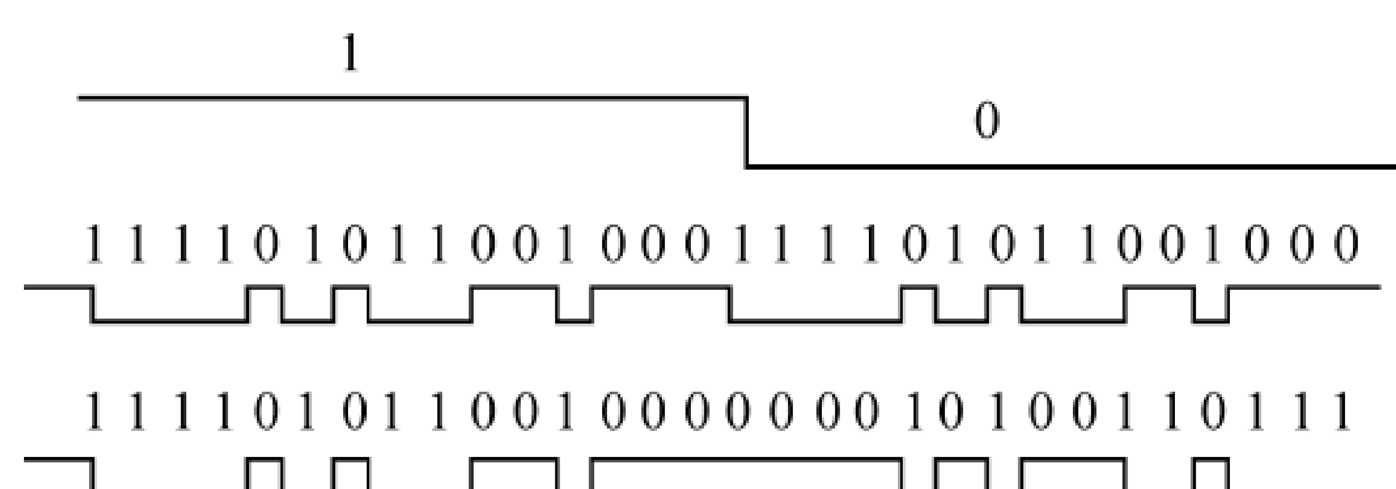



图 1.34　CDMA 扩频

CDMA 手机所使用的 WCDMA、TDCDMA、CDMAX 等技术，均是理论上可以防止低水平窃听的。

（1）CDMA 采用了扩频技术，可以使其信号强度比 GSM 小得多。

（2）CDMA 的呼叫都使用相同的频率，大量的 CDMA 信号共用一个频谱，大大增加了监听分析的难度。

（3）CDMA 手机各自的信号带有不同的随机码，原数据信号的带宽被扩展后经载波调制发射出去。在接收端则使用完全相同的高速伪随机码，在最后环节才将接收的宽带信号还原成窄带信号（解扩）来实现通信的。随机码的使用也大大增加了监听分析的难度。

但是，所有这些并不能保证其不可窃听，只不过窃听难度要比 GSM 难得多。

### 1.2.7 NFC 泄露

#### 1. 概述

近场通信（Near Field Communication，NFC）是一种由非接触式射频识别（RFID）及其互连技术演变而来的短距高频的无线电技术，最先由飞利浦半导体（现恩智浦半导体公司）、诺基亚和索尼共同研制开发。

该技术规范定义了两个 NFC 设备之间基于 13.56MHz 频率的无线通信方式。使用 NFC，不需要卡，也不需要读卡器，只要有两台 NFC 设备，就可以在 20cm 距离内进行数据交换，传输速度有 106kbps、212kbps 或者 424kbps 三种。

从 2006 年诺基亚公司推出第一部 NFC 手机开始，其后陆续有不少设备加入了 NFC 功能。目前，NFC Forum 在全球拥有数百个成员，包括索尼、飞利浦、LG、摩托罗拉、NXP、NEC、三星、英特尔、中国移动、华为、中兴等公司。

#### 2. NFC 在手机上的应用

NFC 技术在手机上应用主要有以下 5 类。

（1）接触通过（Touch and Go），如门禁管理、车票和门票等，用户将存储车票证或门控密码的设备靠近读卡器即可，也可用于物流管理。

（2）接触支付（Touch and Pay），如非接触式移动支付，用户将设备靠近嵌有 NFC 模块的 POS 机可进行支付，并确认交易。

（3）接触连接（Touch and Connect），如把两个 NFC 设备相连接，进行点对点（Peer-to-Peer）数据传输，例如下载音乐、图片互传和交换通讯录等。

（4）接触浏览（Touch and Explore），用户可将 NFC 手机接靠近街头有 NFC 功能的智能公用电话或海报，来浏览交通信息等。

（5）下载接触（Load and Touch），用户可通过 GPRS 网络接收或下载信息，用于支付或门禁等功能，如前述，用户可发送特定格式的短信至家政服务员的手机来控制家政服务员进出住宅的权限。

也就是说，可以凭借一部手机走遍天下。

#### 3. NFC 数据泄露

近距离通信的方便性，也给犯罪者以有机可乘。现在已经有犯罪分子近距离获取别人银行卡信息的报道，所以有关媒体提醒人们不要将银行卡等与手机放在一起，也要提防别人近距离刷走自己重要卡中的信息。

## 1.3 系统扫描型攻击

网络拓扑结构、网络地址、端口开放状况、运行什么样的操作系统、存在哪些漏洞以及用户口令等关系到信息系统的运行和安全状态，它们都可以称为信息系统敏感数据。

获取与网络有关的敏感数据的手段是网络扫描，也称为网络信息采集。内容包括地址扫描、端口扫描和漏洞扫描。网络扫描是网络管理人员进行网络维护常用的手段，也是攻击者进行攻击踩点、确定攻击目标和攻击方法的基本手段。

获取用户口令的手段是口令破解。攻击者获得了用户口令，就可以冒充合法身份进入被攻击系统。

### 1.3.1 网络扫描

网络扫描的目的是判断某个 IP 地址上有无活动主机或某台主机是否在线、到达该目标的路由以及有关端口的打开状况。

**1. 地址扫描**

地址扫描可以直接使用操作系统的有关命令进行，下面是几种常用的命令。

1）ping 命令

ping 是潜水艇人员的专用术语，表示回应的声纳脉冲。在网络中，用 ping 命令向目标主机发送 ICMP 回显请求报文，并等待 ICMP 回显应答，从而检测网络的连通情况和分析网络速度。

ping 命令的完整格式如下。

```
ping [-t] [-a] [-n count] [-l size] [-f] [-i ttl] [-v tos] [-r count] [-s count] [-j computer-list] | [-k computer-list] [-w timeout] destination-list
```

各参数的含义如下。

-t：不断 ping 目标主机，直至中断。

-a：以 IP 地址格式来显示目标主机的网络地址。

-n count：指定要 ping 的次数，默认值为 4。

-l size：指定发送到目标主机的数据包的大小，默认为 32B，最大值是 65527B。

-f：在数据包中发送“不要分段”标志，数据包就不会被路由上的网关分段。

-i ttl：将“生存时间”字段设置为 ttl 指定的值。

-v tos：将“服务类型”字段设置为 tos 指定的值。

-r count：在“记录路由”字段中记录传出和返回数据包的路由。count 可以指定最少 1 台，最多 9 台计算机。

-s count：指定 count 指定的跃点数的时间戳。

-j computer-list：利用 computer-list 指定的计算机列表路由数据包。连续计算机可以被中间网关分隔（路由稀疏源）的 IP 允许的最大数量为 9。

-k computer-list：利用 computer-list 指定的计算机列表路由数据包。连续计算机不能被中间网关分隔（路由严格源）的 IP 允许的最大数量为 9。

-w timeout：指定超时间隔，单位为毫秒。

destination-list：指定要 ping 的远程计算机。

2）tracert 命令

tracert 是一个路由跟踪程序，它通过向目标发送不同的 IP 生存时间（TTL）值的 ICMP 回显数据包来确定到目标所选择的路由。tracert 程序工作时，首先发送一个 TTL=1（ms） 的回显数据包，以后每次递增 1；它要求每个路由器在转发数据包之前先将 TTL 减 1，当 TTL=0 时，路由器将返回 ICMP Time Exceeded。这样，直到目标响应或 TTL 达到最大值后，可以根据每次到达的路由接口列表得到到达目标的路由信息。它的主要选项如下。

-d：防止将中间路由器解析为它们的 IP 地址。

-h maximum_hops：指定搜索到目标地址的最大跳跃数，默认值为 30。

-j host_list：是一系列用空格分隔的带点十进制 IP 地址，用来表示一系列由一个或多个路由器隔开的中间目标。最大数量为 9，仅在 IPv4 中才使用。

-w timeout：指定超时时间间隔，程序默认的时间单位是毫秒。默认值为 4000。

-4：强制 tracert 使用 IPv4。

-6：强制 tracert 使用 IPv6。

target_name：目标主机的名称或 IP 地址。

3）pathping 命令

pathping 命令是一个路由跟踪工具，它将 ping 和 tracert 命令的功能与这两个工具所不提供的其他信息结合起来。pathping 命令在一段时间内将数据包发送到将到达最终目标的路径上的每个路由器，然后根据从每个跃点返回的数据包计算结果。由于命令显示数据包在任何给定路由器或链接上丢失的程度，因此可以很容易地确定可能导致网络问题的路由器或链接。它的主要选项如下。

-n hostnames：不将地址解析成主机名。

-h maximum hops：搜索目标的最大跃点数。

-g host-list：沿着主机列表释放源路由。

-p period：在 ping 之间等待的毫秒数。

-q num_queries：每个跃点的查询数。

-w time-out：每次等待回复的毫秒数。

-i 地址：使用指定的源地址。

-4 IPv4：强制 pathping 使用 IPv4。

-6 IPv6：强制 pathping 使用 IPv6。

4）who 命令

who 命令主要用于查看当前在线上的用户情况。它的主要选项如下。

-a：显示所有用户的所有信息。

-m：显示运行该程序的用户名，和 who am I 的作用一样。

-q：只显示用户的登录账号和登录用户的数量，该选项优先级高于其他任何选项。

-u：在登录用户后面显示该用户最后一次对系统进行操作距今的时间。

-h：显示列标题。

5）ruser 命令

ruser 是一个 UNIX 命令，可以生成登录到远程机的用户列表。它的主要选项如下。

-a：即使没有用户登录也提供报告。

-h：按主机名的字母顺序排序。

-i：按空闲时间排序。

-l：提供类似于 who 命令的更长的清单。

-u：按用户数量排序。

6）finger 命令

finger（端口 79）是一个 UNIX 命令，用于提供站点及用户的基本信息。它的主要选项如下。

-s：显示用户注册名、实际姓名、终端名称、写状态、停滞时间和登录时间等信息。

-l：除了用-s 选项显示的信息外，还显示用户主目录、登录 Shell、邮件状态等信息，以及用户主目录下的.plan、.project 和.forward 文件的内容。

-p：除了不显示.plan 文件和.project 文件以外，与-l 选项相同。

7）host 命令

host 是一个 UNIX 命令，可以把一个主机名解析到一个网络地址或把一个网络地址解析到一个主机名，得到很多信息，包括操作系统、计算机和网络的很多数据。它的必要选项如下。

-a：等同于-v-t。

-C：在需要认证的域名服务器上查找 SOA 记录。

-d：等同于-v。

-l：列出一个域内所有的主机。

-i：反向查找。

-N：改变点数。

-r：不使用递归处理。

-R：指定 UDP 包数。

-T：支持 TCP/IP 模式。

-v：运行时显示详细的处理信息。

-w：永远等待回复。

-W：指定等待回复的时间。

-4：用于 IPv4 的查询。

-6：用于 IPv6 的查询。

8）netstat

该命令可以使用户了解到自己的主机是怎样与因特网相连接的。它的主要选项如下。

-r：显示本机路由表的内容。

-s：显示每个协议（包括 TCP、UDP 和 IP）的使用状态。

-n：以数字表格形式显示地址和端口。

-a：显示所有主机的端口号。

### 2. 端口扫描

在 TCP/IP 网络中，端口号是主机上提供的服务的标识。例如，FTP 服务的端口号为 21，Telnet 服务的端口号为 23，DNS 服务的端口号为 53，HTTP 服务的端口号为 80 等。入侵者知道了被攻击主机的地址后，还需要知道通信程序的端口号。一个打开的端口就是一个潜在的入侵通道。只要扫描到相应的端口已打开，就知道目标主机上运行着什么服务，以便采取针对这些服务的攻击手段。下面介绍几种常用的端口扫描技术。

1）全连接扫描与半连接扫描

TCP 连接通过三次握手（three-way handshake）建立。图 1.35 表示了一个建立 TCP 连接的三次握手过程。若主机 B 运行一个服务器进程，则它要首先发出一个被动打开命令，要求它的 TCP 准备接收客户进程的连接请求，然后服务器进程就处于“听”状态，不断检测有无客户进程发起连接的请求。

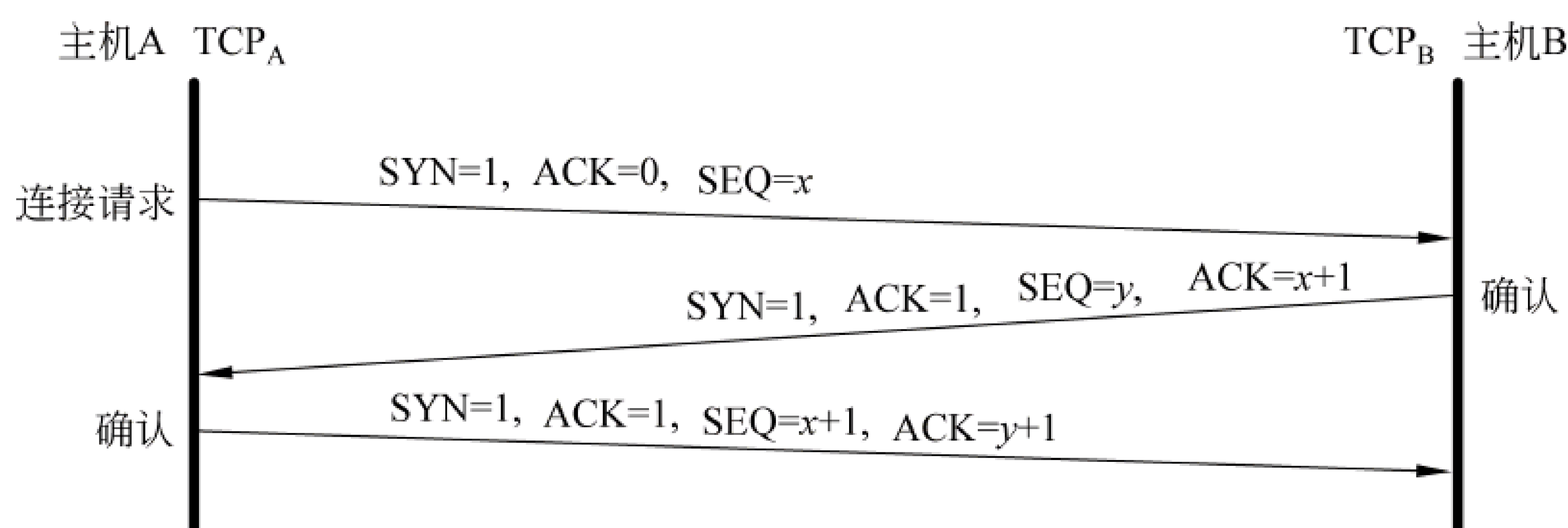



图 1.35 建立 TCP 连接的三次握手过程

（1）若主机 A 中运行有客户进程，当它需要服务器的服务时，就要向它的 TCP 发出主动连接请求：用 SYN=1 和 ACK=0 表示连接请求，用 SEQ=$x$ 表示选择了一个序号。主机 B 收到 A 的连接请求报文，就完成了第一次握手。

（2）主机 B 如果同意连接，其 TCP 就向 A 发回确认报文：用 SYN=1 和 ACK=1 表示同意连接，用 ACK=$x$+1 表示对 $x$ 的确认，用 SEQ=$y$ 表示 B 选择的一个序号。主机 A 接收到该确认报文，完成第二次握手。

（3）接着，主机 A 的 TCP 就还要向主机 B 发出确认：用 SYN=1 和 ACK=1 表示同意连接，用 ACK=$y$+1 表示对 $y$ 的确认，同时发送 A 的第一个数据 SEQ=$x$+1。主机 B 收到主机 A 的确认报文，完成第三次握手过程。

完成这样一个三次握手，才算建立了可靠的 TCP 连接，才能可靠地传输数据报文，也可以获取端口是否开放的信息。这种扫描称为全连接扫描或 TCP connect 扫描。但是，这种扫描往往会被远程系统记入日志。为避免被记入日志，可以使用半开放扫描——TCP SYN 扫描。因为，当客户端发出一个 SYN 连接请求报文后，如果收到了远程目标主机的 ACK/SYN 确认，就说明远程主机的该端口是打开的；而若没有收到远程目标主机的 ACK/SYN 确认，而是收到 RST 数据报文（表明连接出现了问题），就说明远程主机的该端口没有打开。这样

对于扫描要获得的信息已经足够了，也不会在目标主机的日志中留下记录。这种扫描称为半连接扫描或 SYN 扫描。

2）TCP FIN 扫描

FIN 是释放连接的数据报文，表明发送方已经没有数据要发送了。很多日志不记录这类报文。TCP FIN 扫描的原理是向目标端口发送 FIN 报文，当 FIN 数据包到达一个关闭的端口时，会返回一个 RST 的回复；当 FIN 数据包到达一个开放的端口时，该包将被忽略，没有回复。由此可以判断一个端口是关闭还是打开的。这种方法还可以用来区别操作系统是 Windows，还是 UNIX。

这种方法比 SYN 扫描更隐蔽，也被称为秘密扫描。但是，有的系统不管端口打开与否，一律回复 RST。这时，FIN 扫描就不适用了。

3）认证扫描

前面介绍的扫描方法有一个共同特点：判断一个主机中哪个端口上有进程在监听。认证扫描的特点与之不同，它是利用认证协议，获取运行在某个端口上进程的用户名（userid）。其基本思路是：尝试与一个 TCP 端口建立连接，如果连接成功，扫描器发送认证请求到目的主机的 TCP 端口 113。

认证扫描同时也被称为反向认证扫描，因为即使最初的 RFC 建议的是一种帮助服务器认证客户端的协议，然而在实际的实现中也考虑了反向应用（即客户端认证服务器）。

4）UDP ICMP 端口不能到达扫描

这种方法与上面几种方法的不同之处在于使用的是 UDP。由于这个协议很简单，所以扫描变得相对比较困难。这是由于打开的端口对扫描探测并不发送一个确认，关闭的端口也并不需要发送一个错误数据包。不过，许多主机在一个未打开的 UDP 端口上收到一个数据包时，会返回一个 ICMP_PORT_UNREACH 错误，由此可以判断被扫描端口是否关闭的。UDP 和 ICMP 错误都不保证能到达。因此采用这种扫描还必须实现在一个包看上去是丢失的时候能重新传输机制。这种扫描方法是很慢的，并且需要具有 root 权限。

5）UDP recvfrom()和 write()扫描

当非 root 用户不能直接读到端口不能到达错误时，Linux 能间接地在它们到达时通知用户。例如，对一个关闭的端口的第 2 个 write()调用将失败。在非阻塞的 UDP 套接字上调用 recvfrom()时，如果 ICMP 出错还没有到达时会返回 EAGAIN（重试）。ICMP 到达时返回 ECONNREFUSED（连接被拒绝）。这也是用来查看端口是否打开的一项技术。

6）乱序扫描

乱序扫描就是对扫描的端口号集合随机地产生扫描顺序，并且每次的扫描顺序不同。这就给入侵检测系统的发觉端口扫描带来困难。

### 3. 网络扫描工具

1）NMap

网址：http://www.insecure.org/nmap。

NMap 是运行在 Linux/UNIX 下的一个功能非常强大的扫描工具，被称为扫描之王。它支持多种协议（如 TCP、UDP、ICMP 等）扫描，可以用来查看有哪些主机以及其上运行何种服务。

NMap 的扫描方式如下。

（1）TCP connect 扫描。

（2）TCP SYN (half open) 扫描。

（3）TCP FIN、Xmas 或 NULL (stealth) 扫描。

（4）TCP ftp proxy (bounce attack) 扫描。

（5）使用 IP 分片包的 SYN/FIN 扫描。

（6）TCP ACK 和 Window 扫描。

（7）UDP raw ICMP port unreachable 扫描。

（8）ICMP ping 扫描。

（9）TCP ping 扫描。

（10）Direct (non portmapper) RPC 扫描。

（11）通过 TCP/IP 堆栈探测远程主机操作系统和 Reverse-ident 扫描等。

2）SuperScan

SuperScan（见图 1.36）是一款 Windows 平台上的具有 TCP connect 端口扫描、ping 和域名解析等功能的工具，能较容易地做到对指定范围内的 IP 地址进行 ping 和端口扫描。

SuperScan 的功能如下。

（1）通过 ping 来检验 IP 是否在线。

（2）IP 和域名相互转换。

（3）检验目标计算机提供的服务类别。

（4）检验一定范围的目标计算机是否在线和端口情况。

（5）工具自定义列表检验目标计算机是否在线和端口情况。

（6）自定义要检验的端口，并可以保存为端口列表文件。

（7）SuperScan 自带一个木马端口列表 trojans.lst，通过这个列表可以检测目标计算机是否有木马，也可以自己定义修改这个木马端口列表。

3）Wireshark

Wireshark（见图 1.37）是一个网络封包分析软件，其功能是截取流经本地网卡的数据流量进而对其进行分析。通常的应用包括网络管理员用来解决网络问题，网络安全工程师用来检测安全隐患，开发人员用来测试协议执行情况，读者用来学习网络协议。

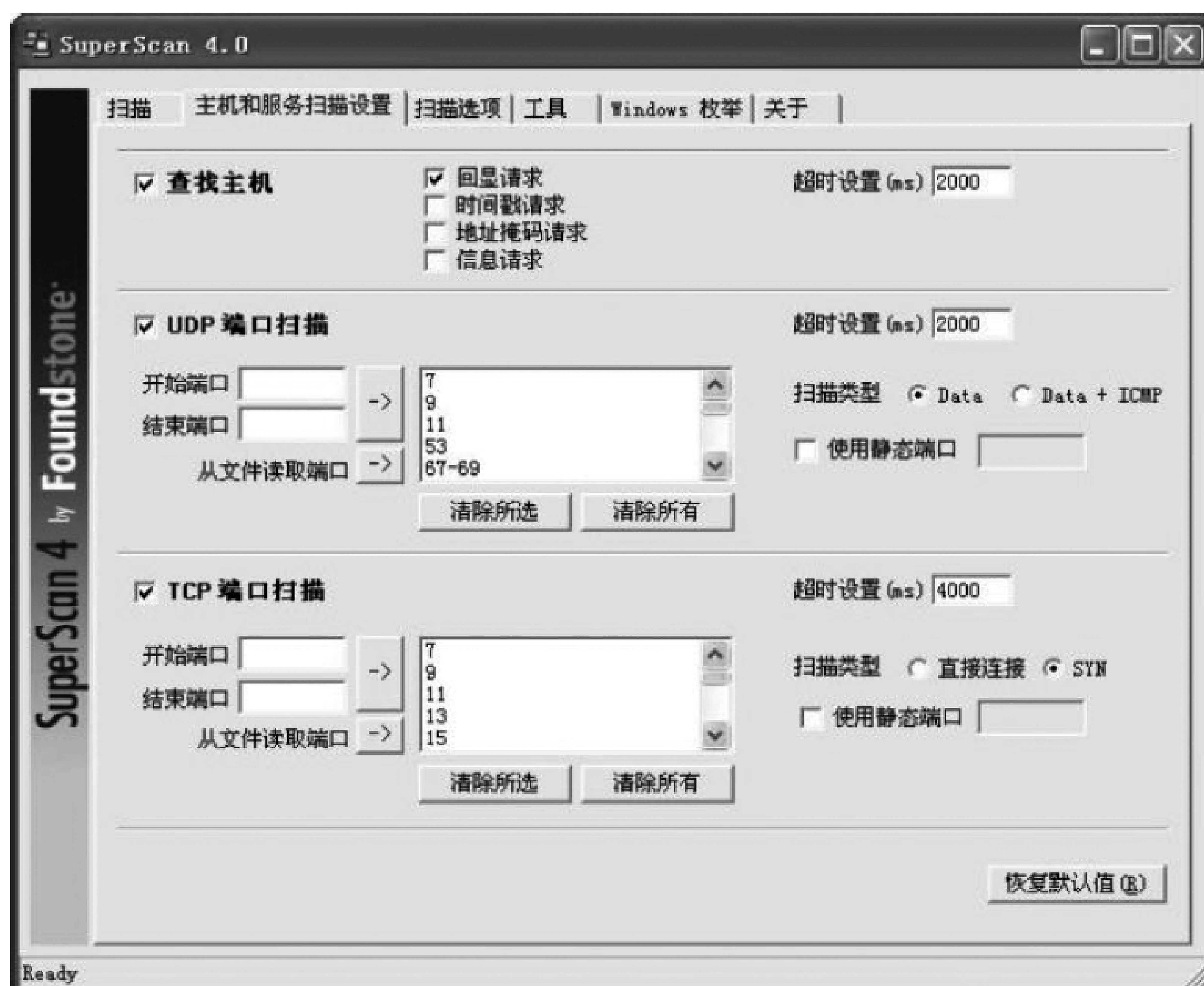



图 1.36　SuperScan 主界面




图 1.37　Wireshark 主界面

图形界面的 Wireshark 使用十分便捷，选取监听的网卡之后，主界面中会显示所有的数据流量。双击任意条目，则可以根据协议的层次拆分该数据流。Wireshark 内置了基本的网络协议，可以方便地查询包括但不局限于 IP、TCP、UDP、HTTP、FTP 和 SMB 等常见的协议内容。

### 1.3.2 漏洞扫描

系统安全漏洞也称为系统脆弱性（vulnerability），是系统缺陷和不足。管理人员可以通过漏洞扫描对所管理的系统和网络进行安全审计，检测系统中的安全脆弱环节，所以也称为网络安全扫描。攻击者则可以通过漏洞扫描找到入侵攻击的缺口实施攻击。

常言道："苍蝇不叮无缝的蛋。"对于信息系统的攻击基本上都是利用系统的漏洞进行的。非法用户可利用系统安全漏洞获得计算机系统的额外权限，在未经授权的情况下访问或提高其访问权，危害计算机系统的正常运行。

攻击者扫描到系统的漏洞，测试出目标主机的漏洞信息后，往往会先通过使用插件（功能模块技术）进行模拟攻击，或者采用漏洞库的匹配方法，制定出攻击的策略。网络管理者也会针对这些漏洞制定相关对策。

#### 1. 系统安全漏洞的类型

对于漏洞可以从不同的角度进行分类，来讨论它们的特点。

1）基于触发主动性的漏洞分类

（1）主动触发漏洞。该漏洞可以被攻击者直接用于攻击，如直接访问他人计算机。

（2）被动触发漏洞。这种漏洞必须有计算机操作人员配合才能起作用。例如，攻击者给某人发一封带有特殊的 jpg 图片文件的邮件，接收者只有打开该图片文件，才会导致某个漏洞被触发，使系统被攻击；若接收者不看这个图片，则不会受攻击。

2）基于发现时间的漏洞分类

（1）已发现很久的漏洞。厂商发布补丁或修补方法已经有一段时间，广为知晓。由于很多人已经进行了修补，因此宏观危害较小。

（2）刚发现的漏洞。厂商刚发布补丁或修补方法，知道的人还不多。这种漏洞相对于已发现很久的漏洞危害性较大，若此时使用蠕虫或傻瓜化程序，就会导致大批系统受到攻击。

（3）0day 漏洞。还没有公开，或因私下交易而形成的漏洞。这类漏洞会导致目标受到精确攻击，危害非常大。

3）基于系统或部位的漏洞分类

（1）操作系统漏洞。指计算机操作系统本身所存在的问题或技术缺陷。操作系统产品提供商通常会定期对已知漏洞发布补丁程序提供修复服务。

（2）Web 服务器漏洞。主要包括物理路径泄露、CGI 源代码泄露、执行任意命令、缓冲区溢出、拒绝服务、SQL 注入、条件竞争和跨站脚本执行漏洞。

（3）不同服务相互感染漏洞。有时候在一台服务器上会运行多种网络服务，如 Web 服务、FTP 服务等。这就很可能会造成服务之间的相互感染，攻击者只要攻击一种服务，就可以利用相关的技术作为平台，攻陷另一种服务。

（4）数据库服务器漏洞。如某些数据库服务器在处理请求数据时存在缓冲区溢出漏洞，

远程攻击者可能利用此漏洞控制服务器，向数据库服务器发送畸形请求触发漏洞，最终导致执行任意指令。

（5）应用程序漏洞。这种漏洞由应用程序编写时的错误导致。目前，大部分应用程序都有数百万行代码。但人们只需要几分钟就可以开启后门或者安放“定时炸弹”。

（6）内存覆盖漏洞。内存覆盖漏洞主要为内存单元可指定，写入内容可指定。这样就能执行攻击者想执行的代码（如缓冲区溢出漏洞、格式化字符串漏洞、PTrace 漏洞、Windows 2000 的硬件调试寄存器用户可写漏洞）或直接修改内存中的机密数据。

4）基于成因的漏洞分类

（1）操作性漏洞。可以分为两种情形。

① 写入内容被控制。导致可伪造文件内容、权限提升或直接修改重要数据（如修改存贷数据）。

② 内容信息被输出。包含内容被打印到屏幕、记录到可读的日志文件、产生可读的 core 文件等。

（2）配置漏洞。可以分为如下两种。

① 系统配置漏洞。多源于管理员疏漏，如共享文件配置漏洞、服务器参数配置漏洞、使用默认参数配置的漏洞等。

② 网络结构配置漏洞。多与网络拓扑结构有关，如将重要设备与一般设备设置在同一网段等。

（3）协议漏洞。这种漏洞主要源于 Internet 上的现行协议在设计之初仅考虑了效率和可靠性，没有考虑安全性。这类漏洞很多。后面将要介绍的 ARP 欺骗、IP 源地址欺骗、路由欺骗、TCP 会话劫持、DNS 欺骗和 Web 欺骗等都是由于协议漏洞引起的。此外还有 UDP Flood（循环）攻击（基于 UDP 端口漏洞）、SYN Flood 攻击、Land 攻击、Smurt 攻击、WinNuke 攻击、Fraggle 攻击和 Ping to death 攻击等都源于相关协议漏洞。

（4）程序漏洞。程序漏洞缘于程序设计的复杂性、程序设计语言的漏洞和运行环境的不可预见性。下面是一些常见程序漏洞。

① 缓冲区溢出漏洞。

② 格式字符串漏洞。

③ BIND 漏洞。

④ Finger 漏洞。

⑤ SendMail 漏洞。

### 2. 常用漏洞扫描器

目前已经开发出了大量的扫描器。下面仅列举几例。

1）ISS/SAFESuite（应用层风险评估工具）

ISS（Internet Security Scanner）始于 1992 年，最初由 Christopher Klaus 发布，是一个小小的开放源代码扫描器，但功能强大，不过价格也昂贵。

它可以用来检查使用 TCP/IP 网络连接的主机是否会受到攻击，可以扫描以下漏洞：

① 一些默认的包头，如是否存在 guest、bbs 等。
② IP 包头。
③ Decode Alias。
④ Sendmail。
⑤ 匿名 FTP。
⑥ NIS。
⑦ NFS。
⑧ rusers。

SAFESuite 是 ISS 的最新版本，功能更强，效率更高，不仅可以运行在 UNIX 下，还可以运行在 Windows 下。它不仅能广泛检查各种服务，还能对所发现的漏洞提供如下信息。

① 位置。
② 有关描述。
③ 正确的应对建议。

2）Nessus

Nessus（见图 1.38）是一款可以运行在 Linux、BSD、Solaris 以及其他一些系统上的远程漏洞扫描与分析软件，它采用 B/S 架构的方式安装，以网页的形式向用户展现。用户登录之后可以指定对本机或者其他可访问的服务器进行漏洞扫描。Nessus 的扫描程序与漏洞库相互独立，因而可以方便地更新其漏洞库，同时提供多种插件的扩展和一种语言 NASL（Nessus Attack Scripting Language）用来编写测试选项，极大地方便了漏洞数据的维护、更新。在扫描完成后，Nessus 还可以生成详尽的用户报告，包括脆弱性、漏洞修补方法以及危害级别等，以方便后续加固工作。

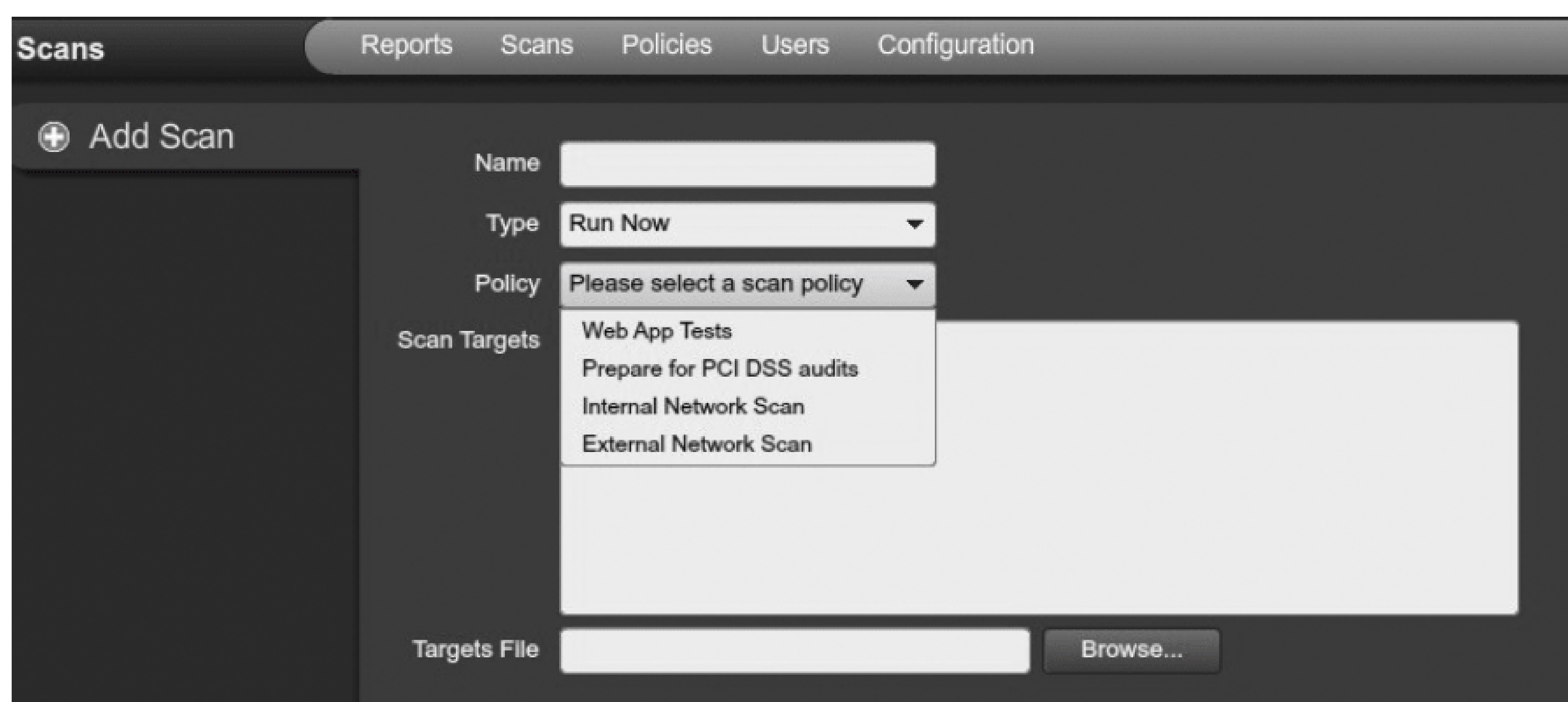



图 1.38 Nessus 主界面

3）Mysfind

Mysfind 是著名的扫描器 pfind 的加强版，主要用于扫描 Printer 漏洞和 Unicode 漏洞。Printer 漏洞可以让攻击者取得系统的控制权，Unicode 可以让攻击者随意操作系统内的文件甚至完全控制系统。它采用多线程扫描系统漏洞，速度快，结果准。

Mysfind 是一个命令行程序，格式如下：

```
sfind 漏洞类型 开始 IP 地址 结束 IP 地址
```

它有 3 种扫描方式。

-all：扫描所有漏洞。

-e：扫描 Printer 漏洞，可以让攻击者取得系统的控制权。

-u：扫描 Unicode 漏洞，可以让攻击者随意操作计算机内的文件甚至完全控制系统。

扫描结束以后，结果自动保存在 sfind.txt 文件中。

4）X-Scan

X-Scan 能够扫描大范围网段中存在漏洞的主机。它采用多线程方式对指定 IP 地址（或单机）进行安全漏洞检测，支持插件功能，可以在图形和命令两种界面下操作，扫描内容包括远程操作系统类型及版本、标准端口状态、端口 BANNER 信息、SNMP 信息、CGI 漏洞、IIS 漏洞、RPC 漏洞、SQL-Server、FTP-Server、SMTP-Server、POP3-Server、NT-Server 和注册表信息等。

5）Zenoss

Zenoss 是商业服务器监控工具 Zenoss Enterprise 的一个开源版本，全部由 Python 语言编写。它支持 Nagios plugin format（Nagios 插件格式），所以许多 Nagios 的插件也可以用于 Zenoss。Zenoss 的一个突出的地方是它强大而又容易使用的用户接口。

6）AppScan

AppScan 是 IBM 公司推出的一款 Web 应用安全测试工具，它采用黑盒测试的方式，可以扫描常见的 Web 应用安全漏洞。

7）Nikto

Nikto 是一款非常全面的 Web 扫描器，能在 200 多种服务器上扫描出 2000 多种有潜在危险的文件、CGI 及其他问题。它也使用 LibWhiske 库，但通常比 Whisker 更新得更为频繁。

8）N-Stealth

N-Stealth 是 ZMT 公司出品的一款商业的 Web 站点安全扫描软件，同时也有可以免费使用的版本，只是功能没有商业版本的多，漏洞库也不支持自动更新。

## 实验 2　系统扫描

### 1. 实验目的

（1）了解扫描攻击的基本原理。

（2）掌握常用扫描工具的基本用法。

（3）学习扫描器程序设计的基本方法。

### 2. 实验内容

（1）使用两种扫描器软件进行扫描，包括 SATAN、流光、CIS 和 SuperScan 等。对扫

描结果进行统计分析，并提出被扫描系统的安全改进方案。

（2）比较两种扫描器的功能、特点和效果。

（3）演示自己设计的端口或漏洞扫描程序，并记录演示的扫描过程及结果。

（4）建立漏洞库。

（5）运行自己设计的、基于漏洞库的、高效率的扫描软件，用它进行端口和漏洞扫描，并进行主机脆弱性分析。

**3. 实验准备**

（1）设计实验的环境。

（2）比较几种常用的扫描器，选定一两种实用扫描器。

（3）设计使用扫描器的步骤。

（4）设计漏洞库建立的方法和步骤。

（5）设计一个可以演示扫描过程和结果的扫描器程序。

（6）设计一个基于漏洞库设计并实现一个高效率的扫描软件，可以进行端口和漏洞扫描，并可以进行主机脆弱性分析。

**4. 推荐的分析讨论内容**

（1）分析网络扫描器在网络管理和网络安全方面的作用。

（2）其他发现或想到的问题。

### 1.3.3 口令破解

口令机制是资源访问的第一道关口。攻破了这道关口，就打开了进入系统的第一道大门。所以口令攻击是入侵者最常用的攻击手段。口令攻击可以从破解口令和屏蔽口令保护两个方面进行。下面主要介绍口令破解技术。

**1. 口令破解的基本技术**

口令破解首先要获取口令文件，然后采取一定的攻击技术进行口令的破解。下面介绍口令破解的基本方法。

1）口令字典猜测破解法

攻击者基于某些知识，编写出口令字典，然后对字典进行穷举或猜测攻击。表 1.5 为口令字典的构造方法。

目前，Internet 上已经提供了一些口令字典，从一万到几十万条，可以下载。此外，还有一些可以生成口令字典的程序。利用口令字典可以通过猜测方式进行口令破解攻击。

2）穷举破解法

有人认为使用足够长的口令或者使用足够完善的加密模式，就不会被攻破。事实上没有攻不破的口令，这只是个时间问题。如果有速度足够快的计算机能尝试字母、数字、特殊字符的所有组合，将最终能破解所有的口令。这种类型的攻击方式通过穷举口令空间获

表 1.5　口令字典的构造方法

| 序号 | 口 令 类 型 | 实　　例 | 序号 | 口 令 类 型 | 实　　例 |
|---|---|---|---|---|---|
| 1 | 规范单词 | computer | 19 | 医药词汇 | vitamin |
| 2 | 反写规范单词 | retupmoc | 20 | 技术词汇 | Ruter |
| 3 | 词首正规大写 | Computer | 21 | 商品 | beer |
| 4 | 反拼写与反大写 | computeR | 22 | 用户标识符 | woodc |
| 5 | 缩写 | TCP | 23 | 反写用户标识符 | cdoow |
| 6 | 带点缩写 | T.C.P | 24 | 串接用户标识符 | woodc-woodc |
| 7 | 缩写后带点 | TCP. | 25 | 截短用户标识符 | woo |
| 8 | 略写 | etc. | 26 | 串接用户标识符并截短 | woodcwood |
| 9 | 专有名词缩写，带点 | Ph.D | 27 | 单字符构成串 | bbbbbb |
| 10 | 专有名词缩写，不全大写 | kHz | 28 | 键盘字母 | asdfgh |
| 11 | 姓 | Bush | 29 | 文化名人 | Beethoven |
| 12 | 名 | Tom | 30 | 年月日 | 040723 |
| 13 | 所有格 | Bob　s | 31 | 电话号码 | 5863583 |
| 14 | 动词变化 | see, sees, saw, seen | 32 | 邮政编码 | 214036 |
| 15 | 复数 | books | 33 | 证件号码 | 20010612345 |
| 16 | 法律用语 | legal | 34 | 门牌号码 | AB3579 |
| 17 | 地名（城/街/山/河等） | BeiJing | 35 | 车牌号码 | 苏-w12345 |
| 18 | 生物词汇 | Dog | ⋮ | ⋮ | ⋮ |

得用户口令，称为穷举破解法或蛮力破解，也称为强行攻击。例如先从字母 a 开始，尝试 aa、ab、ac、…，然后尝试 aaa、aab、aac、…。

3）组合破解法

词典破解法只能发现词典单词口令，但是速度快。穷举破解法能发现所有的口令，但是破解时间很长。鉴于很多管理员要求用户使用字母和数字，用户的对策是在口令后面添加几个数字，如把口令 computer 变成 computer99。使用强行破解法又非常费时间。由于实际的口令常常很弱（可以通过对字典或常用字符列表进行搜索或经过简单置换而发现的口令），这时可以基于词典单词而在单词尾部串接几个字母和数字，这就是组合破解法。

4）其他破解类型

① 社会工程学：通过对目标系统的人员进行游说、欺骗、利诱，获得口令或其部分。

② 偷窥：观察别人输入口令。

③ 搜索垃圾箱。

**2. 口令破解工具**

1）Cain & Abel（穷人的 L0phtCrack）

网址：http://www.oxid.it/cain.html。

类别：免费。

平台：Windows。

简介：Cain & Abel 是一个针对 Windows 操作系统的免费口令恢复工具。它通过如下多种方式轻松地实现口令恢复：网络嗅探、破解加密口令（使用字典或强行攻击）、解码被打乱的口令、显示口令框、显示缓存口令和分析路由协议等。该工具的源代码不公开。

2）DSniff（一流的网络审计和渗透测试工具）

网址：http://naughty.monkey.org/~dugsong/dsniff/。

类别：开放源码。

平台：Linux/BSD/UNIX/Windows。

简介：DSniff 是由 Dug Song 开发的一套包含多个工具的软件套件。其中，dsniff、filesnarf、mailsnarf、msgsnarf、rlsnarf 和 webspy 可以用于监视网络上的数据（如口令、E-mail、文件等），arpspoof、dnsspoof 和 macof 能很容易地载取到攻击者通常难以获取的网络信息（如二层交换数据），sshmitm 和 webmitm 则能用于实现重写 SSH 和 HTTPS 会话达到 monkey-in-the-middle 攻击。在 http://www.datanerds.net/~mike/dsniff.html 可以找到 Windows 平台上的移植版。

3）John the Ripper（格外强大、灵活、快速的多平台哈希口令破解器）

网址：http://www.openwall.com/john/。

类别：开放源码。

平台：Linux/BSD/UNIX/Windows。

简介：John the Ripper 是一个快速的口令破解器，支持多种操作系统，如 UNIX、DOS、Win32、BeOS 和 OpenVMS 等。它设计的主要目的是用于检查 UNIX 系统的弱口令，支持几乎所有 UNIX 平台上经 crypt 函数加密后的口令哈希码，也支持 Kerberos AFS 和 Windows NT/2000/XP LM 哈希码等。

4）L0phtCrack 4（Windows 口令审计和恢复程序）

网址：http://www.atstake.com/research/lc/。

类别：商业。

平台：Linux/BSD/UNIX/Windows。

简介：L0phtCrack 从独立的 Windows NT/2000 工作站、网络服务器、主域控制器或 Active Directory 上正当获取或者从线路上嗅探到的加密哈希值里破解出 Windows 口令，含有词典攻击、组合攻击、强行攻击等多种口令猜解方法。

5）网络刺客

网络刺客是一个强大的网络安全工具，扫描只是其中的一个功能。它的扫描功能包括共享扫描、端口扫描和口令扫描猜测等。

## 1.4　欺骗型攻击

在网络环境下，通信主体之间的认证也是基于数字进行的。这种非直接的认证为欺骗（spoofing）提供了机会。广义地说，网络欺骗泛指在网络环境下，攻击者冒充已经建立了信任关系的对象中的一方，对另一方进行欺骗，获取有用资源的行为。一般说来，网络欺骗是针对网络协议漏洞实施的。本节讨论几种常见的欺骗型攻击的原理和手段。

### 1.4.1　ARP 欺骗——交换网络监听

在 1.2.5 节中介绍了共享网络中的窃听问题。但是，现在以太网已经从共享进入到交换时代。交换式网络不是共享网络，交换式设备可以准确地将数据报文发给目的主机。这时，在交换网络中，能够直接收听的是群发帧和广播帧，无法直接实施广泛监听。

在这种环境下，由于一个局域网上发往其他网络的数据帧的目标地址都是指向网关的，因此实施监听的一个简单的方法是将安装有 Sniffer 软件的计算机伪装成为网关。这就是地址解析协议（Address Resolution Protocol，ARP）欺骗窃听。

#### 1．ARP 欺骗窃听原理

ARP 是一个将 32 位的 IP 地址翻译成 48 位 MAC 地址的协议。图 1.39 是 ARP 的请求和应答分组格式。

<table>
<tr><td>以太网目的地址(6B)</td><td>以太网源地址(6B)</td><td>帧类型(2B)</td><td>ARP首部(8B)</td><td>源MAC地址(6B)</td><td>源IP地址(4B)</td><td>目的MAC地址(6B)</td><td>目的IP地址(4B)</td></tr>
<tr><td colspan="3">以太网首部</td><td colspan="5">以太网ARP字段</td></tr>
</table>

图 1.39　ARP 请求和应答分组格式

ARP 协议是一个无状态的协议，一旦收到 ARP 应答报文，就会对其高速缓存中的 IP 地址到 MAC 地址的映射记录进行更新，而不会关心之前是否发出过 ARP 请求。ARP 欺骗的核心就是向目标主机发送一个包含伪造的 IP-MAC 映射信息的 ARP 应答报文。当目的主机收到此应答报文后就会更新其 ARP 高速缓存，从而使目标主机将报文发送给错误的对象。这种攻击也称为中间人攻击。

所谓中间人攻击，就是使进行监听的主机插入到被监听主机与其他网络主机之间，利用 ARP 欺骗进行攻击，造成进行监听的主机成为被监听主机与其他网络主机通信的中继。如图 1.40 所示，当主机 A 要给主机 B 发送 IP 包时，在包头中需要填写 B 的 IP 为目标地址，并且这个 IP 包在以太网上传输的时候，还需要进行一次以太包的封装，即填入 B 的 MAC 地址。但是 A 是不知道 B 的 MAC 地址的。为了获得 B 的 MAC 地址，A 就广播一个 ARP 请求包，请求包中填有 B 的 IP 地址，以太网中的所有计算机都会接收这个请求，而正常的情况下只有 B 会给出 ARP 应答包，包中就填充上了 B 的 MAC 地址，并回复给 A。A 得到 ARP 应答后，将 B 的 MAC 地址放入本机缓存，便于下次使用，或者更新已有的 ARP 缓存。

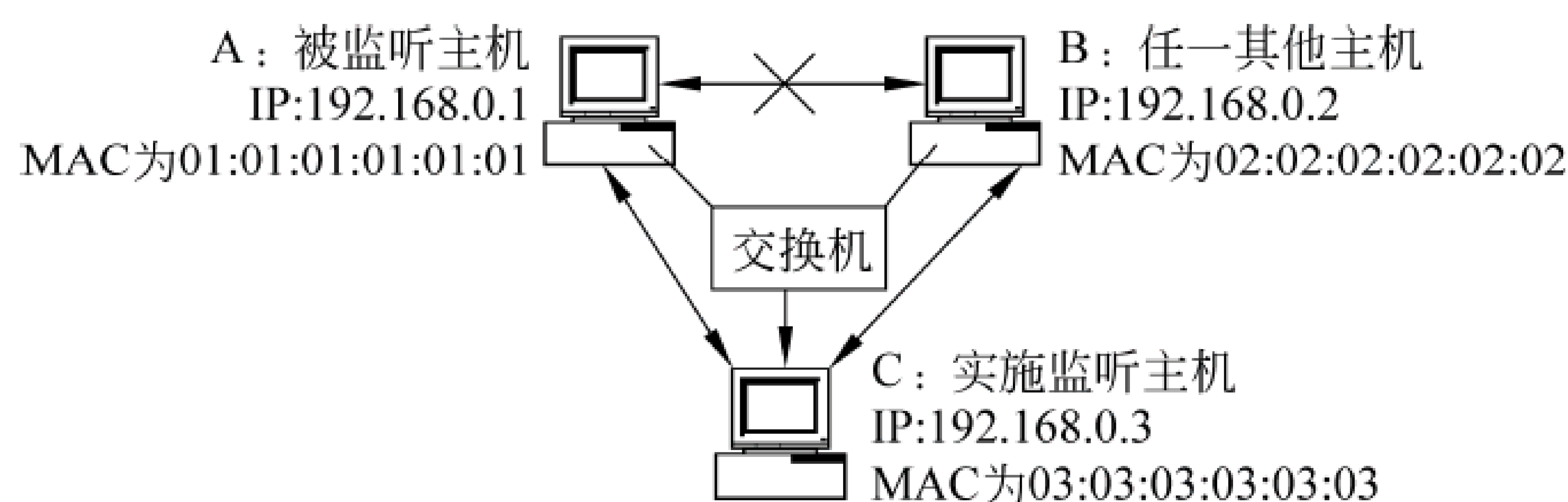



图 1.40　ARP 欺骗窃听

由于 ARP 并不只在发送了 ARP 请求后才接收 ARP 应答，这就会使入侵者有机可乘。例如，局域网中的主机 C 可能会冒充主机 B 向 A 发送一个伪造的 ARP 应答，应答中的 IP 地址为 B 的 IP 地址，而 MAC 地址是主机 C 的 MAC 地址（也可以是另外的主机 D 的 MAC 地址），则当 A 接收到 C 伪造的 ARP 应答后，就会更新本地的 ARP 缓存，这样 A 就会把发向 B 的数据包发送到同一物理网的主机 C（或 D）。这样，C 就是插入的 A 和 B 之间的进行攻击的主机。

**2．ARP 欺骗窃听防范**

1）采用静态 ARP

指定静态 ARP，即将 IP 地址与 MAC 地址绑定。大多数 UNIX 系统支持 ARP 读取指定的 IP 和 MAC 地址对应文件，首先编辑内容为 IP 和 MAC 地址对照的文件，然后使用命令 arp -f /path/to/ipandmacmapfile 读取文件，这样就指定了静态的 ARP 地址，即使接收到 ARP 应答，也不会更新自己的 ARP 缓存，从而使 ARP 欺骗丧失作用。

Windows 系统没有-f 这个参数，但有-s 参数，可以用命令行指定 IP 和 MAC 地址对照关系，如 arp -s 192.168.1.33 00-90-6d-f2-24-00。但除了 Windows XP 外，其他版本的 Windows 平台即使这样做，当接收到伪造的 ARP 应答后，依然会更新自己的 ARP 缓存，用新的 MAC 地址替换掉旧的 MAC 地址，所以无法对抗 ARP 欺骗。而且采用静态 ARP 有一个缺憾，就是如果网络很大的话，工作量会非常大。

2）ARP 监听检测

首先，借助检测 IP 地址和 MAC 地址对应的工具，如 arpwatch，安装了 arpwatch 的系统在发生 MAC 地址变化时会在系统的日志文件中看到如下提示：

```
Apr 21 23:05:00 192.168.1.35 arpwatch:flip flop 192.168.1.33 0:90:6d:f2:24:0 (8:0:20:c8:fe: 15)
Apr 21 23:05:02 192.168.1.35 arpwatch:flip flop 192.168.1.33 8:0:20:c8:fe:15 (0:90:6d:f2: 24:0)
Apr 21 23:05:03 192.168.1.35 arpwatch:flip flop 192.168.1.33 0:90:6d:f2:24:0 (8:0:20:c8:fe: 15)
```

从提示中可以看出，arpwatch 检测到了网关 MAC 地址发生了改变。

其次，借助一些入侵检测系统，如 Snort，也可以起到的一定的检测作用。在 Snort 的配置文件中打开 arpspoof 的 preprocessor 开关并进行配置即可。

3）数据加密

数据加密可以使攻击者即使窃听到，也无法了解内容。

### 3. 监听器

监听器（sniffer）是一种用于捕获网络报文的软件，可以用来进行网络流量分析，找出网络中潜在问题，确定在通信所使用的多个协议中属于不同协议的流量大小，哪台主机承担主要协议的通信，哪台主机是主要的通信目的地，报文发送的时间是多少，主机间报文传送的时间间隔等，是网络管理员的一种常用工具。

监听器只是接收数据，而不向外发送数据，从而能悄无声息地监听到所有局域网内的数据通信，其潜在危害性也在于此。

下面介绍几种常用的监听器。

1）Sniffer Pro

Sniffer Pro 是 NAI 公司开发的一种图形界面嗅探器。它功能强大，能全面监视所有网络信息流量，识别和解决网络问题，是目前唯一能够为七层 OSI 网络模型提供全面性能管理的工具。

2）Libpcap/Winpcap

Libpcap 是 Packet Capture Library（数据包捕获函数库）的缩写与重组。它不是一个监听器，但是它提供的 C 语言函数接口可用于对经过网络接口的数据包的捕获，以支持监听器产品的开发。Winpcap 是 Libpcap 的 Win32 版本。

3）Dsniff

Dsniff 是 Dug Song 编写的一个功能强大的工具软件包，它可以支持多种协议类型，包括 FTP、Telnet、rlogin、Ldap、SMTP、POP、IMAP、IRC、ICQ、MS-CHAP、Npster、Citrix、ICA、PCAnywher、SNMP、OSPF、PPTP、X11、NFS、RIP、VRRP、Microsoft SQL Protocol 等。

4）Tcpdump/Windump

Tcpdump 是一个传统的嗅探器，通过将网卡设置为混杂模式截取帧进行工作。

### 4. ARP 监听软件 arpspoof

arpspoof 是 Dsniff 中的一个组件，是一个基于 ARP 理论的网络监听程序，它的工作原理是这样的：发起 ARP 欺骗的主机向目标主机发送伪造的 ARP 应答包，骗取目标系统更新 ARP 表，将目标系统的网关的 MAC 地址修改为发起 ARP 欺骗攻击的主机 MAC 地址，使数据包都经由发起 ARP 欺骗的主机。这样即使系统连接在交换机上，也不会影响对数据包的截取，由此就轻松地通过交换机实现了网络监听。例如，主机 A 和 B 连接在交换机的同一个 VLAN（虚拟局域网）上。

A 主机的 IP 地址为 192.168.1.37。

B 主机的 IP 地址为 192.168.1.35，MAC 地址为 08-00-20-c8-fe-15。

网关的 IP 地址为 192.168.1.33，MAC 地址为 00-90-6d-f2-24-00。

（1）在 A 主机上看看 A 主机的 ARP 表：

```
C:>arp -a
nbsp; Interface:192.168.1.37
Internet Address Physical Address Type
192.168.1.33 00-90-6d-f2-24-00 dynamic
```

可以看到 A 主机中保留着网关的 IP 地址 192.168.1.33 和对应的 MAC 地址 00-90-6d-f2-24-00。

（2）在 B 主机上执行 arpspoof，将目标指向 A 主机，宣称自己为网关，如下：

```
HOSTB# arpspoof -t 192.168.1.37 192.168.1.33
8:0:20:c8:fe:15 0:50:ba:1a:f:c0 0806 42:arp reply 192.168.1.33 is-at 8:0:20:c8:fe:15
8:0:20:c8:fe:15 0:50:ba:1a:f:c0 0806 42:arp reply 192.168.1.33 is-at 8:0:20:c8:fe:15
8:0:20:c8:fe:15 0:50:ba:1a:f:c0 0806 42:arp reply 192.168.1.33 is-at 8:0:20:c8:fe:15
8:0:20:c8:fe:15 0:50:ba:1a:f:c0 0806 42:arp reply 192.168.1.33 is-at 8:0:20:c8:fe:15
8:0:20:c8:fe:15 0:50:ba:1a:f:c0 0806 42:arp reply 192.168.1.33 is-at 8:0:20:c8:fe:15
8:0:20:c8:fe:15 0:50:ba:1a:f:c0 0806 42:arp reply 192.168.1.33 is-at 8:0:20:c8:fe:15
8:0:20:c8:fe:15 0:50:ba:1a:f:c0 0806 42:arp reply 192.168.1.33 is-at 8:0:20:c8:fe:15
8:0:20:c8:fe:15 0:50:ba:1a:f:c0 0806 42:arp reply 192.168.1.33 is-at 8:0:20:c8:fe:15
8:0:20:c8:fe:15 0:50:ba:1a:f:c0 0806 42:arp reply 192.168.1.33 is-at 8:0:20:c8:fe:15
```

可以看到 B 主机持续向 A 主机发送 ARP 回应包，宣称网关 192.168.1.33 的 MAC 地址是 B 主机自己。此时，在 A 主机上看看 ARP 表的内容：

```
C:>arp -a
Interface: 192.168.1.37
Internet Address Physical Address Type
192.168.1.33 08-00-20-c8-fe-15 dynamic
```

显然 A 主机的 ARP 表已经改变了，网关的 MAC 地址被更新为 B 主机的 MAC 地址。这样，当有数据包发送时，A 主机理所当然地会发到其 ARP 表中网关对应的 MAC 地址 08-00-20-c8-fe-15。可是，A 主机却不知道它的数据实际上发送到了一台别有用心的 B 主机。

（3）但是还不能这样结束。为了让 A 主机不会有明显的感觉，B 主机还必须打开数据转发。

① 在 Linux 中可以使用 sysctl -w net.ipv4.ip_forward=1。

② 在 BSD 系统可以使用 sysctl -w net.inet.ip.forwarding=1。

③ 在 Solaris 系统可以使用 ndd -set /dev/ip ip_forwarding=1。

除了这样打开内核的支持外，也可以选用外部的 fragrouter 等转发软件，如此，就能确保 A 主机能正常工作了。

## 实验 3　监听器工具的使用

### 1. 实验目的

（1）了解监听器的基本原理。

（2）掌握一种监听器工具的基本用法。

### 2. 实验内容

（1）下载并安装一种监听器工具。
（2）使用安装的监听器工具进行网络信息收集。

### 3. 实验准备

（1）设计实验的环境。
（2）选定一种监听器工具，记录其下载网址。
（3）罗列出使用该监听器工具的步骤和方法。

### 4. 推荐的分析讨论内容

（1）分析监听器工具在网络管理和网络安全方面的作用。
（2）其他发现或想到的问题。

## 1.4.2 IP 源地址欺骗

IP 有一个缺陷：它只依据 IP 头中的目的地址发送数据包，而不对数据包中的 IP 地址进行认证。这个缺陷使任何人不经授权就可以伪造 IP 包的源地址。IP 源地址欺骗就是基于这一点，使攻击者可以假冒他人的 IP 地址向某一台主机发送数据包，进行攻击。

攻击者使用 IP 地址欺骗的目的主要有两种。

（1）隐藏自身，对目标主机发送不正常包，使之无法正常工作。

（2）伪装成被目标主机信任的友好主机得到非授权的服务。

### 1. IP 源地址欺骗攻击的基本过程

IP 源地址欺骗是冒用别的主机的 IP 地址用于欺骗第三者。假定有两台主机 S（设 IP 地址为 201.15.192.11）和 T（设 IP 地址为 201.15.192.22），并且它们之间已经建立了信任关系。入侵者 X 要对 T 进行 IP 欺骗攻击，就可以假冒 S 与 T 进行通信。

（1）确认攻击目标。施行 IP 源地址欺骗的第一步是确认攻击目标。下面是容易受到电子欺骗攻击的服务类型。

① 运行 Sun RPC（Sun Remote Procedure Call，Sun 远程过程调用）的网络设备。

② 基于 IP 地址认证的任何网络服务。

③ 提供 R 系列服务的计算机，如提供 rlogin、rsh、rcp 等服务的计算机。

其他没有这类服务的系统所受到的 IP 欺骗攻击虽然也有，但要少得多。

（2）使被冒充的主机无法响应目标主机的会话。当 X 要对 T 实施 IP 源地址欺骗攻击时，就要假冒 S（称为被利用者）与目标主机 T 进行通信。但是，X 并不是真正的 S，而 T 只向 S 回送应答包。这样，就有可能使 S 对 T 的报文产生反应，而将 X 暴露。X 避免自己暴露的办法是让 S 瘫痪，使之无法响应目标主机 T 的数据包。

使 S 瘫痪的办法是对其实施拒绝服务攻击，例如，通过 SYN Flood 攻击使之连接请求被占满，暂时无法处理进入的其他连接请求。通常，黑客会用一个虚假的 IP 地址（可能该

合法 IP 地址的服务器没有开机）向目标主机 TCP 端口发送大量的 SYN 请求。受攻击的服务器则会向该虚假的 IP 地址发送响应。自然得不到回应，得到的是该服务器不可到达的消息。而目标主机的 TCP 会认为这是暂时的不通，于是继续尝试连接，直到确信无法连接。不过这已经为黑客进行攻击提供了充足的时间。

（3）精确地猜测来自目标请求的正确序列数。

X 为了使自己的攻击不露馅的另一个措施是取得被攻击目标 T 主机的信任。由于 TCP 是可靠传输协议，每台主机要对自己发送的所有字节分配序列编号，供接收端确认并据此进行报文装配。在通过三次握手建立 TCP 连接的过程中，客户端首先要向服务器发送序列号 $x$，服务器收到后通过确认要向客户端送回期待的序列号 $x$+1 和自己的序列号。由于序列号的存在，给 IP 欺骗攻击增加了不少难度，要求攻击者 X 必须能够精确地猜测出来自目标机的序列号，否则也会露馅。

那么，如何精确地猜测来自目标机的序列号呢？这就需要知道 TCP 序列号的编排规律。

初始的 TCP 序列号是由 tcp_init( )函数确定的，是一个随机数，并且它每秒钟增加 128 000。这表明，在没有连接的情况下，TCP 的序列号每 9.32 h 会复位一次；而有连接时，每次连接把 TCP 序列号增加 64 000。

随机的初始序列号的产生也是有一定规律的。在 Berkeley 系统中，初始序列号由一个常量每秒钟加 1 产生。

所以，TCP 序列号的估计也并非绝对不可能。但是，除此之外，攻击者还需要估计他的服务器与可信服务器之间的往返时间（RTT）。RTT 一般是通过多次统计平均计算出来的。在没有连接的情况下，TCP 序列号为 128000 • RTT；如果目标服务器刚刚建立过一个连接，就还要加上 64 000。

上述分析是一种理论上的分析。黑客通常的做法是通过对目标主机的合法连接来获得目标主机发送 IP 数据包的序列记录。具体步骤如下。

① 请求连接目标主机。

② 目标主机送回带序列号的回应。

③ 记录序列号并断开连接。

在一般情况下，通过对所记录的序列号的分析，可以猜测出认证要求序列号的规则。

（4）冒充受信主机连接到目标主机。

（5）根据猜出的序列号，向目标主机发送回应 IP 包。

（6）进行系列会话。

**2．IP 源地址欺骗的防范**

IP 源地址欺骗攻击比较普遍，而且产生的危害性很大。下面是 IP 欺骗的一些预防策略。

（1）放弃基于 IP 地址的信任策略。IP 欺骗是基于 IP 地址信任的。而 IP 地址很容易伪造。因此，阻止这类攻击的一种非常简单的方法是放弃以 IP 地址为基础的验证。

（2）使用随机化的初始序列号。序列号是接收方 TCP 进行合法检查的一个重要依据。黑客攻击能够得逞的一个重要因素就是序列号有一定的选择和增加规律。堵塞这一漏洞的方法就是让黑客无法计算或猜测出序列号。Bellovin 提出了一个公式：

$$ISN=M+F（localhost,localport,remotehost,remoteport）$$

其中，$M$ 为 4 μs 定时器，$F$ 为加密 Hash 函数，localhost 为本地主机，localport 为本地端口，remotehost 为远方主机，remoteport 为远方端口。Bellovin 建议 $F$ 是一个结合连接标识符和特殊矢量（随机数，基于启动时间的密码）的 Hash 函数，它产生的序列号不能通过计算或猜测得出。

（3）在路由器中加上一些附加条件。这些条件包括：不允许声称是内部包的外部包（源地址和目标地址都是本地域地址）进入，以防止外部攻击者假冒内部主机的 IP 欺骗；禁止带有内部资源地址的内部包出去，以防止内部用户对外部站点的攻击。

（4）配置服务器，降低 IP 欺骗的可能：分析自己的服务器，看哪些服务容易遭受 IP 欺骗攻击，并考虑这些服务有无保留的必要。

（5）使用防火墙和其他抗 IP 欺骗的产品。例如，防火墙决定是否允许外部的 IP 数据包进入局域网，对来自外部的 IP 数据包进行检验。假如来自外部的数据包声称有内部地址，它一定是欺骗包。如果数据包的 IP 地址不是防火墙内的任何子网，它就不能离开防火墙。

### 1.4.3 路由欺骗

在 TCP/IP 网络中，IP 包的传输路径完全由路由表决定。因此，如果攻击者能控制路由表，则可以控制自己发送的 IP 包到达希望的目标，进行监听或其他攻击。下面介绍两种路由欺骗方法。

#### 1. IP 源路由欺骗

IP 报文首部具有“源站选路”可选项，可以指定到达目的站点的路由。在正常情况下，若目的主机有应答或其他信息回送源站点，就可以按照源站所选路由逆向路径回复。

假定有两台主机 S（设 IP 地址为 201.15.192.11）和 T（设 IP 地址为 201.15.192.22），之间已经建立了信任关系。若有攻击者 X 想冒充 S 从 T 处获得某种服务，则它可以按照下面的步骤进行。

（1）攻击者 X 进行 IP 地址欺骗，将自己发往 T 的源地址修改为 201.15.192.11，目标地址写为 201.15.192.22。

（2）将路由表写为 X 到 T 的路由（例如 X 所在局域网的路由器 GX）。

这样，T 要发送到 S 的应答，实际上按照 X 指定的路由的逆向路径，送回到 X 所在局域网的路由器 GX。X 通过监听收取该数据包。当然，要求路由器 GX 所在的局域网中不包括 S。

防范 IP 源路由欺骗的主要方法是关闭主机和路由器上的源路由选项。此外，也可以通过路由器配置，阻挡那些来自外部却声称是内部主机的报文。

#### 2. RIP 路由欺骗

路由信息协议（Routing Information Protocol，RIP）是一种基于选择算法的内部或域内路由选择协议。距离向量算法（DVA）的路由选定原则是，到一个目的站的最少“路由中继”（hop，跳）数或到那个目的站路径的费用。为适应网络的变化，要求域内路由器之间

动态地（每 30s）交换信息，内容包括每个路由器可以到达哪些网络，这些网络有多远等。信息交换用 IP 数据包进行，并用 UDP 作为传输协议。

这样，若攻击者在网上发布假的路由信息，声称路由器 GX 可以使 T 发往 S 的数据包最快到达，再通过 ICMP 重定向来欺骗服务器路由器和主机，将 T 到 S 的正常路由器标志为失效，就可以诱使 T 将数据包发到攻击者控制的路由器 GX。

### 1.4.4 TCP 会话劫持

会话劫持（session hijack）就是攻击者作为第三者，隐秘地加入到他人的会话中，发送恶意数据，或者对他人的会话进行监听，甚至接替其中一方与另一方会话。在网络中进行会话劫持往往是利用了通信协议的漏洞，通过嗅探与欺骗两种手段结合进行。

#### 1. TCP 会话劫持的基本原理

（1）TCP 使用端到端的连接，即 TCP 用（源 IP，源 TCP 端口号，目的 IP，目的 TCP 端号）来唯一标识每一条已经建立连接的 TCP 链路。

（2）TCP 在进行数据传输时，其首部有两个非常重要的字段：序号（seq）和确认序号（ackseq）。

① 序号指出了本报文中传送的数据在发送主机所要传送的整个数据流中的顺序号。

② 确认序号指出了发送本报文的主机希望接收的对方主机中下一个八位组的顺序号。

对于一台主机来说，其收发的两个相邻 TCP 报文之间的序号和确认序号与所携带 TCP 数据净荷（payload）的多少有数值上的关系：它所要发出的报文中的 seq 值应等于它所刚收到的报文中的 ackseq 的值，而它所要发送的报文中 ackseq 的值应为它所收到的报文中 seq 的值加上该报文中所发送的 TCP 净荷的长度。

（3）在 TCP 连接中，只是刚开始连接时进行一次 IP 地址的验证，在连接过程中 TCP 应用程序只跟踪序列号，而不进行 IP 地址验证。因此，一旦同一网段上的入侵者获悉目标主机的序列号规律，就可以假冒该目标主机的受信机与该目标主机进行通信，把原来目标主机与其受信机之间的会话劫持过去。

#### 2. TCP 会话劫持过程

会话劫持一般采取如下 3 步进行。

（1）找一个同网段的活动会话。会话劫持的第一步要求攻击者找一个位于同一网段的活动会话。这要求攻击者嗅探在子网上的通信。攻击者将寻找诸如 FTP 之类的一个已经建立起来的 TCP 会话。在共享网络中，查找这种会话是很容易的。在交换网络中则需要攻击 ARP 协议。

（2）猜测正确的序列号码。进行 TCP 传输时，传输的数据的每一个字节必须有一个序列号码。这个序列号用来保持跟踪数据和提供可靠性。最初的序列号码是在 TCP 协议握手的第一步生成的。目的地系统使用这个值确认发出的字节。这个序列号字段长度有 32B。这就意味着可能有大约 4 294 967 295 个序列号。

（3）把合法的用户断开。一旦确定了序列号，攻击者就能要把合法用户断开，以免露

出破绽。断开合法用户可以采用拒绝服务攻击、源路由欺骗或者向用户发送一个重置命令。如果这些步骤取得成功，攻击者现在就可以控制这个会话。只要这个会话能够保持下去，攻击者就能够通过身份验证进行访问。

#### 3. 会话劫持攻击工具

1）Juggernaut

Juggernaut 是由 Mike Schiffman 开发的一个可以用来进行 TCP 会话攻击的网络监听器，它是一个开放的自由软件。可以运行于 Linux 操作系统的终端机上，安装和运行都很简单。可以设置值、暗号或标志这 3 种不同的方式来通知 Juggernaut 程序是否对所有的网络流量进行观察。例如，一个典型的标记就是登录暗号。无论何时 Juggernaut 发现这个暗号，就会捕获会话，这意味着黑客可以利用捕获到的用户密码再次进入系统。

2）Hunt

Hunt 是 Kra 开发的一个用来监听、截取和劫持网络上的活动会话的程序。

3）TTY Watcher

TTY Watcher 是一个免费的程序，允许人们监视并且劫持一台单一主机上的连接。

4）IP Watcher

IP Watcher 是一个商用的会话劫持工具，它允许监视会话并且获得积极的反会话劫持方法。它基于 TTY Watcher，此外还提供一些额外的功能，IP Watcher 可以监视整个网络。

### 1.4.5 DNS 欺骗

域名系统（Domain Name System，DNS）是一个将主机域名和 IP 地址互相映射的数据库系统，它的安全性对于互联网的安全有着举足轻重的影响。但是由于 DNS Protocol 在自身设计方面存在缺陷，安全保护和认证机制不健全，造成 DNS 自身存在较多安全隐患，导致其很容易遭受攻击。DNS 欺骗（DNS spoofing）就是利用 DNS 漏洞进行的攻击行为。

#### 1. DNS 的服务过程

设有如图 1.41 所示的 3 台主机。其中，S 向 A 提供 DNS 服务，A 想要访问 B（www.ccc.com）。这个过程如下。

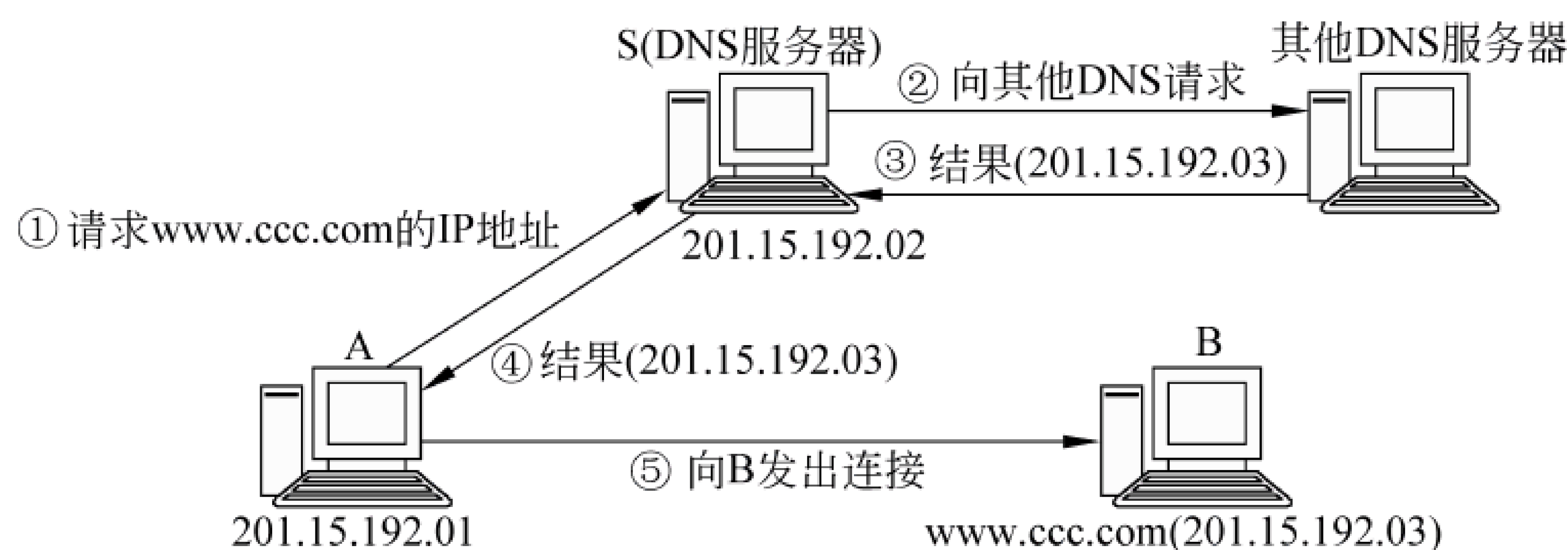



图 1.41　DNS 工作过程示意图

① A 向 S 发一个 DNS 查询请求，要求 S 告诉 www.ccc.com 的 IP 地址，以便与之通信。

② S 查询自己的 DNS 数据库，若找不到 www.ccc.com 的 IP 地址，即向其他 DNS 服务器求援，逐级递交 DNS 请求。

③ 某个 DNS 服务器查到了 www.ccc.com 的 IP 地址，向 S 返回结果。S 将这个结果保存在自己的缓存中。

④ S 把结果告诉 A。

⑤ A 得到了 B 的地址，就可以访问 B 了（如向 B 发出连接请求）。

在上述过程中，如果 S 在一定的时间内不能向 A 返回要查找的 IP 地址，就会向 A 返回主机名不存在的错误信息。

**注意：**DNS 客户端的查询请求和 DNS 服务器的应答数据包是依靠 DNS 报文的 ID 标识来相互对应的。这个 ID 是随机产生的。在进行域名解析时，DNS 客户端首先用特定的 ID 号向 DNS 服务器发送域名解析数据包。DNS 服务器找到结果后使用此 ID 给客户端发送应答数据包。DNS 客户端接收到应答包后，将接收到的 ID 与请求包的 ID 对比，如果相同则说明接收到的数据包是自己所需要的，如果不同就丢弃此应答包。

### 2．DNS 欺骗原理

DNS 有两个重要特性。

（1）DNS 对于自己无法解析的域名，会自动向其他 DNS 服务器查询。

（2）为提高效率，DNS 会将所有已经查询到的结果存入高速缓存（cache）。

正是这两个特点，使得 DNS 欺骗（DNS spoofing）成为可能。实施 DNS 欺骗的基本思路是让 DNS 服务器的高速缓存中存有错误的 IP 地址，即在 DNS 高速缓存中放一个伪造的记录。为此，攻击者需要做两件事。

（1）先伪造一个用户的 DNS 请求。

（2）再伪造一个查询应答。

但是，在 DNS 包中还有一个 16 位的查询标识符（Query ID），它将被复制到 DNS 服务器的相应应答中，在多个查询未完成时，用于区分响应。所以，回答信息只有 Query ID 和 IP 都吻合才能被 DNS 服务器接收。因此，进行 DNS 欺骗攻击，还须精确地猜测出 Query ID。由于 Query ID 每次加 1，只要通过第一次向将要欺骗的 DNS 服务器发一个查询包并监听其 Query ID 值，随后再发送设计好的应答包，包内的 Query ID 就是要预测的 Query ID。

### 3．DNS 欺骗过程

下面结合图 1.41，介绍 DNS 欺骗的一次过程。

（1）入侵者先向 S（DNS 服务器）提交查询 www.ccc.com 的 IP 地址请求。

（2）S 向外递交查询请求。

（3）入侵者立即伪造一个应答包，告诉 www.ccc.com 的 IP 地址是 201.15.192.04（往往是入侵者的 IP 地址）。

（4）查询应答被 S（DNS 服务器）记录到高速缓存中。

（5）当 A 向 S 提交查询 www.ccc.com 的 IP 地址请求时，S 将 201.15.192.04 告诉 A。

#### 4．DNS 欺骗的局限性

DNS 欺骗有如下局限性。

（1）入侵者不能替换 DNS 高速缓存中已经存在的记录。

（2）高速缓存中的记录具有一定的生存期，过期就会被刷新。

### 1.4.6　Web 欺骗与钓鱼网站

#### 1．Web 欺骗的作用

Web 欺骗的攻击者会创建整个 WWW 世界的影像副本。用户进入该影像 Web 的入口，实际是进入到攻击者的 Web 服务器。此时，攻击者就可以肆意实施如下攻击。

（1）可将用户的一切活动置于攻击者的监控之下，攻击者就会获得用户 ID、密码、浏览过的网页和停留的时间等。

（2）能以受攻击者的名义将错误或者易于误解的数据发送到真正的 Web 服务器。

（3）以任何 Web 服务器的名义发送数据给受攻击者。

通过这些活动，攻击者可以实施诈骗进行获利。典型的例子是假冒金融机构偷盗客户的信用卡信息。

钓鱼网站（phishing，是 phone 与 fishing 的组合）就是 Web 欺骗技术的应用。随着电子商务的发展，钓鱼网站一直处于快速蔓延和持续增长势态。如图 1.42 所示，金山网络在《2011—2013 中国互联网安全研究报告》中提供的数据表明，2013 年金山毒霸拦截的钓鱼网站数量超过了 240 万个，与 2011 年、2012 年相比，环比分别增长 340%和 36.7%。每日新增拦截钓鱼网站数量超过 6400 个。

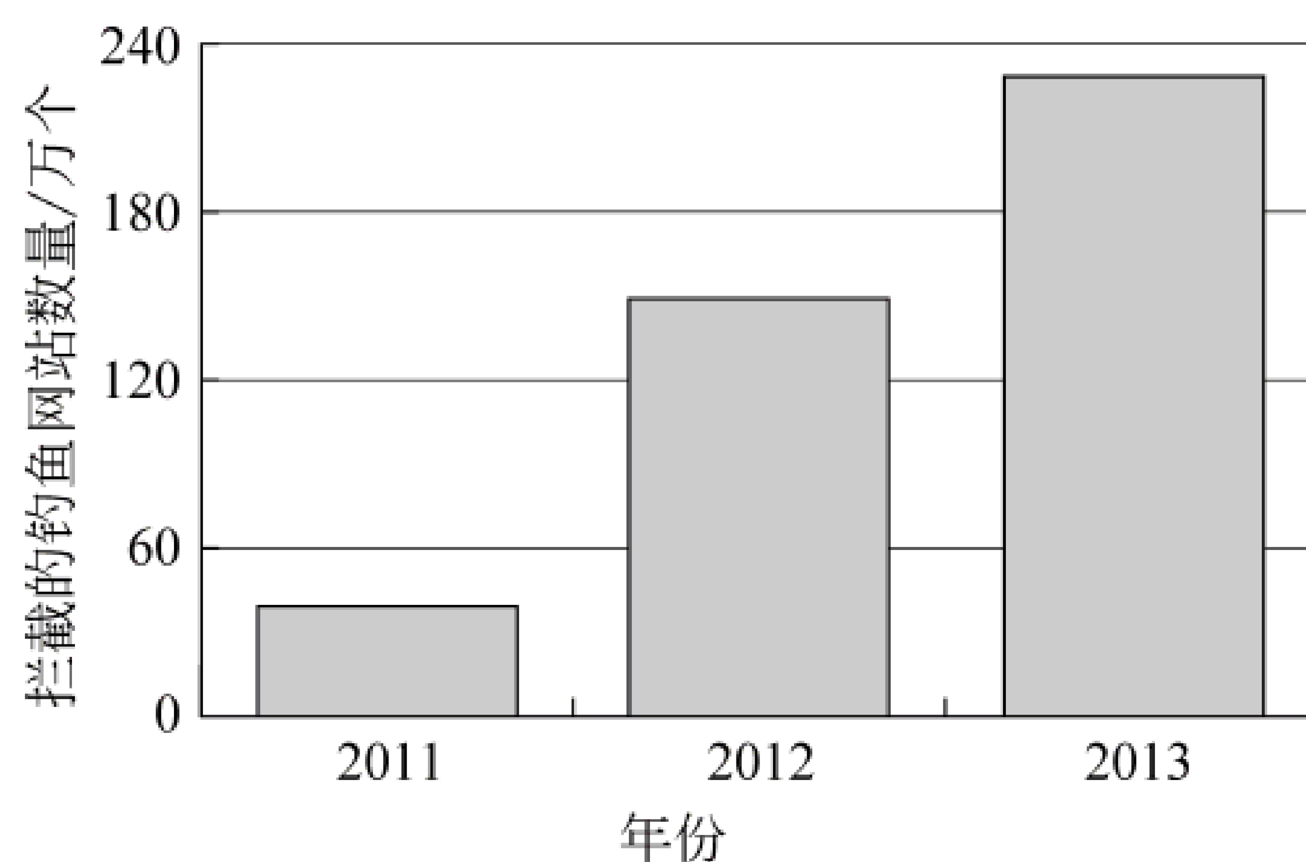



图 1.42　2011—2013 年金山毒霸拦截的钓鱼网站情况

360 安全中心将这些钓鱼网站分为如下 7 种类型。

（1）设置中奖骗局。这类网站冒充游戏网站、QQ、CCTV 等知名栏目、门户网站等发布中奖、奖励信息，用与官方网站极为相似的页面骗取访问者的 QQ、游戏账号密码，或者骗取中奖者支付领取奖品的相关费用。

（2）各种预测网站。这些网站会利用彩民梦想中大奖的心理，进行各种收费预测，网站上会列出很多预测准确的记录来骗取彩民信任，吸引彩民加入会员，购买所谓专家的预测资料，大部分网站还会打出中国福利彩票的官方字样，甚至还有不少是涉及六合彩、赌

球方面的网站。彩民如果相信了他们所谓的预测号码、操纵比赛、预测比赛结果的骗局，往往都会损失惨重。

（3）黑马股票的骗局。这些网站会使用知名证券公司的名称，或者干脆使用一些不存在的证券公司名称来构建网站，网站以保证高额利润为诱饵，向股民推荐涨停股和黑马股，向被骗股民收取高额会员费和保密费，甚至直接让股民投资给他们做代理炒股。受骗股民往往会损失几万元或数十万元。

（4）虚假购物网站。这类网站会在网民购物时出现在买家或卖家的QQ/旺旺上，买卖双方经常发送各种与商品有关的链接，而钓鱼网站就会掺杂其中，页面通常会模仿淘宝、拍拍、支付宝、财付通等与购物有关的网站，骗取账号密码或钱财。

另外还有一些用极低价格来吸引顾客的购物网站，如钻石、手表、手机、充值卡，也是涉及购物欺诈的一种。

（5）仿冒官方网站登录。无论人们在网络上做什么，最常遇到的就是需要输入注册账号，输入用户名、密码。各种模仿官方网站登录的钓鱼网站，仿真度非常高，和官方网站几乎一模一样，即时通信（Instant Message，IM）、网络游戏、邮箱、购物网站、银行，几乎只要是可以登录的网站，就有相应的钓鱼网站。

（6）仿冒的医疗、药品相关网站。近年来，销售假药、劣质药品、假冒医疗机构的现象也时有发生，这类网站也是钓鱼欺诈的重要类型。这些网站的危害不只是骗人钱财，更重要的是会影响人的健康。

（7）假冒的下载网站。这类网站往往有着和官方下载页面相似的页面，但是却提供含有木马的下载链接，这属于用钓鱼的方式来推广木马。

**2．Web 欺骗的基本原理**

Web 欺骗基于如下 3 个基本原理。

（1）目前注册一个域名没有任何要求。利用这一点，攻击者会抢先或特别设计注册一个有欺骗性的站点。

（2）Cookie 欺骗。在浏览器/服务器系统中，Cookie 指由服务器创建的数据文件，这个文件会记录客户在浏览器中所输入的任何文字和选择，包括用户 ID、密码、浏览过的网页、选择的商品等。这些数据被存放到用户计算机硬盘的一个小的文件（C:\\Documents and Settings\\用户名\\Cookies）中。当该用户再光临同一个网站时，Web 服务器会先看看有没有它上次留下的 Cookie 资料，有的话，就会依据 Cookie 里的内容来判断使用者，为该用户送出特定的网页内容。它最早是网景公司的前雇员 Lou Montulli 在 1993 年 3 月发明的。显然，获取客户 Cookie 文件，就可以获取客户的许多敏感信息。另外，Cookie 可以由服务器创建和修改，所以攻击者也可以修改用户的 Cookie 内容。Cookie 欺骗就是在只对用户进行 Cookie 验证的系统中，通过伪造的 Cookie 获得登录权限。

（3）Session 欺骗。在 Web 中，Session 是用来记录浏览器端数据（如用户 ID 等敏感数据）的对象。这些数据存放在服务器端。当一个访问者首次访问一个网页时，服务器就会为其创建一个新的、独立的 Session 对象，为该次会话分配一个会话标识 ID，并把该次会话的会话 ID 的特殊加密版本的 Cookie 发送给客户端。当浏览器关闭时，这个会话 ID 即消失。

所以 Session 生命期仅为一次会话的时间，一般为几十分钟。Session 欺骗就是在只对用户进行 Session 验证的系统中通过伪造的 Session 获得登录权限。

### 3．Web 欺骗的技巧

1）改写 URL

首先，攻击者改写 Web 页中的所有 URL 地址，这样它们指向了攻击者的 Web 服务器而不是真正的 Web 服务器。

2）表单欺骗

在 URL 改写的基础上，表单欺骗将会进行得非常自然。当受攻击者提交表单后，所提交的数据进入了攻击者的服务器。攻击者的服务器能够观察甚至修改所提交的数据。同样，在得到真正的服务器返回信息后，攻击者在将其向受攻击者返回以前也可以为所欲为。

3）设计攻击的导火索

为了开始攻击，攻击者必须以某种方式引诱受攻击者进入攻击者所创造的错误的 Web。黑客往往使用下面若干种方法。

（1）把错误的 Web 链接到一个热门 Web 站点上。

（2）如果受攻击者使用基于 Web 的邮件，可以将它指向错误的 Web。

（3）创建错误的 Web 索引，指示给搜索引擎。

4）完善攻击

前面描述的攻击相当有效，但是它还不是十分完美的。黑客往往还要创造一个可信的环境，包括各类图标、文字、链接等，提供给受攻击者各种各样的可信的暗示，以隐藏一切尾巴。

5）状态信息

状态信息显示在浏览器底部。Web 欺骗中涉及两类信息。

（1）当鼠标指针放置在 Web 链接上时，连接状态显示链接所指的 URL 地址。

（2）当 Web 连接成功时，连接状态将显示所连接的服务器名称。

这两项信息都容易使攻击者露出尾巴——URL 或服务器名称。为此，攻击者往往通过 JavaScript 编程来弥补这两项不足。由于 JavaScript 能够对连接状态进行写操作，而且可以将 JavaScript 操作与特定事件绑定在一起，所以，攻击者完全可以将改写的 URL 状态恢复为改写前的状态。这样 Web 欺骗将更为可信。

### 4. 钓鱼网站的推广方式

（1）即时通信推广。旺旺、MSN 和 QQ 都属于即时通信工具，骗子会利用即时工具跟会员沟通，经常以下列迷惑性情况直接向会员发送钓鱼链接。

① 专柜和银行联合搞特价活动。

② 支付宝被监管。

③ 宝贝拍不了，被监管。

④ 商品出现下架等。

（2）在论坛、博客、微博、问答类网站等发布帖子链接钓鱼网站，往往会用“我知道一个特别好的网站”等推荐方式。

（3）通过手机短信和 E-mail 等批量发布链接，如冒充“银行密码重置邮件”，冒充颁奖等欺骗用户点击进入钓鱼网站。

（4）在网站上制作仿冒 QQ、阿里旺旺等知名软件的弹窗，吸引用户进入钓鱼网站。

（5）在搜索引擎、中小网站投放广告，吸引用户点击钓鱼网站链接，此种手段多被假医药网站、假机票网站所采用。

（6）使用恶意导航网站、恶意下载网站弹出仿真悬浮窗口，点击后进入钓鱼网站。

（7）利用与正规网站极为相似的域名，混淆视听，使用户难判真假。表 1.6 为采用混淆视听方法的几个典型钓鱼网站。

**表 1.6　几个采用混淆视听方法的钓鱼网站**

| 假冒网站域名 | 正规网站域名 | 攻击方式 |
|---|---|---|
| www.1cbc.com.cn | www.icbc.com.cn(中国工商银行) | 金融诈骗 |
| www.1enovo.com | www.lenovo.com(联想公司) | 假冒 |
| www.chsic.com.cn | www.chsi.com.cn(中国高等教育学生信息网) | 发布虚假学历证书信息 |
| www.cnbank-yl.com | www.chinaunionpay.com(中国银联) | 金融诈骗 |
| www.chinacharity.cn.net | www.chinacharity.cn(中华慈善总会) | 利用废弃域名骗取善款 |

金山网络 2013 年发布的数据表明，访问到钓鱼网站的 4 个主要渠道是：搜索引擎（58.7%）、QQ 等聊天工具（30.4%）、手机短信（10.8%）和网页弹窗广告（0.1%）。

### 5. 钓鱼网站的逃避手段

（1）境外注册域名，逃避网络监管。

（2）连续转账操作，迅速转移网银款项。

### 6. 钓鱼网站的防范

1）查验“可信网站”

通过第三方网站身份诚信认证辨别网站的真实性。不少网站已在网站首页安装了第三方网站身份诚信认证——“可信网站”，可帮助网民判断网站的真实性。“可信网站”验证服务通过对企业域名注册信息、网站信息和企业工商登记信息进行严格交互审核来验证网站真实身份，通过认证后，企业网站就进入中国互联网络信息中心（CNNIC）运行的国家最高目录数据库中的“可信网站”子数据库中，从而全面提升企业网站的诚信级别，网民可通过点击网站页面底部的“可信网站”标识确认网站的真实身份。网民在网络交易时应养成查看网站身份信息的使用习惯。企业也要安装第三方身份诚信标识，加强对消费者的保护。

2）核对网站域名

假冒网站一般和真实网站有细微区别，有疑问时要仔细辨别其不同之处，例如在域名方面，假冒网站通常将英文字母 I 替换为数字 1，CCTV 被换成 CCYV 或者 CCTV-VIP 这样

的仿造域名。

3）比较网站内容

假冒网站上的字体样式不一致，并且模糊不清。假冒网站上没有链接，用户可单击栏目或图片中的各个链接看是否能打开。

4）查询网站备案

通过 ICP 备案可以查询网站的基本情况、网站拥有者的情况，对于没有合法备案的非经营性网站或没有取得 ICP 许可证的经营性网站，根据网站性质，将予以罚款，严重的关闭网站。

5）查看安全证书

大型的电子商务网站都应用了可信证书类产品，这类网站网址都是 https 打头的，如果发现不是 https 开头，应谨慎对待。

### 1.4.7　伪基站攻击

#### 1. 基站与伪基站

在移动通信中，为了扩大覆盖范围，除了可以采用卫星外，一种有效的方法是采用图 1.43（a）所示的蜂窝技术。如图 1.43（b）所示，基站是指在一定的无线电覆盖区中，通过移动通信交换中心，与移动电话终端之间进行信息传递的无线电收发信电台。一个小区的用户通信都要通过该小区的基站。或者说，基站可以控制所在区的用户通信，获取该小区用户的通信有关信息。

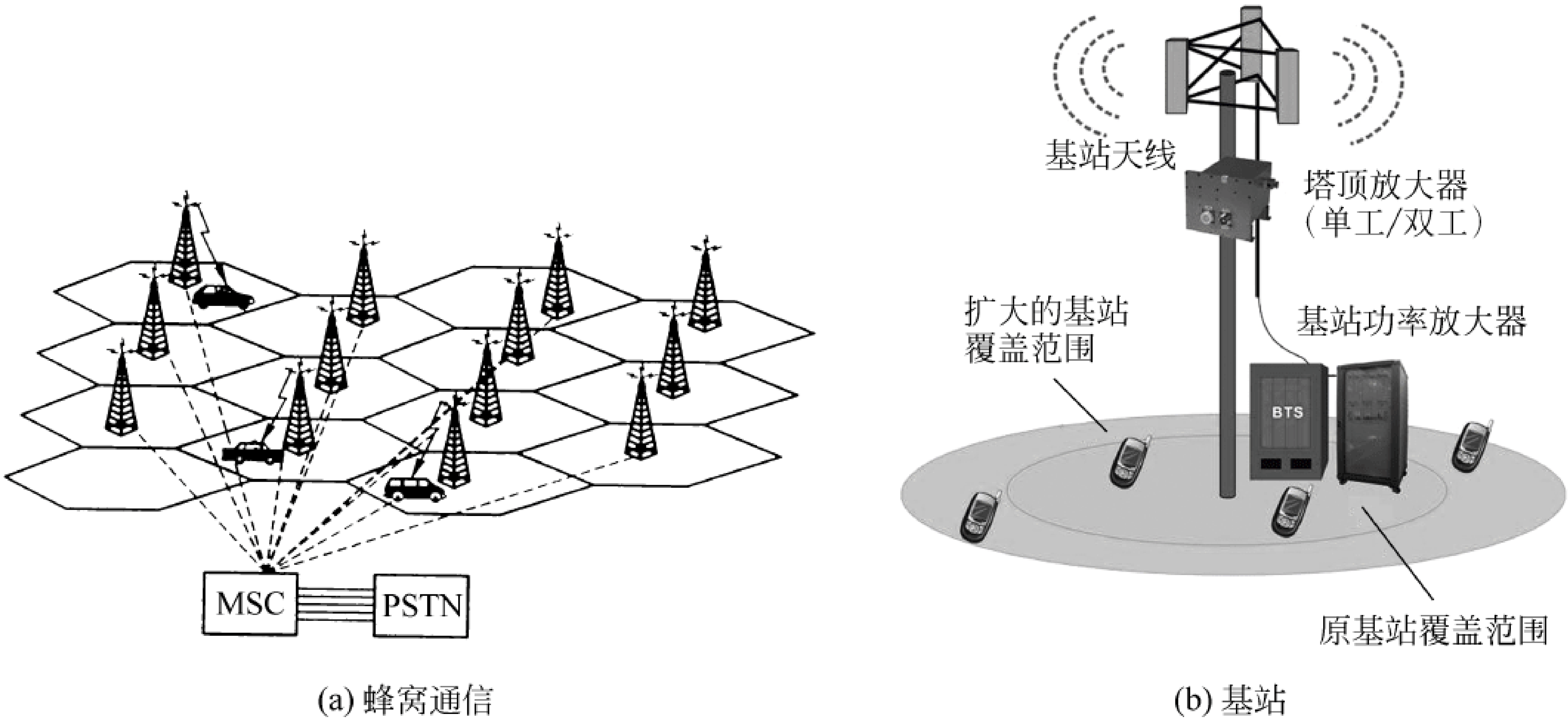



图 1.43　蜂窝通信及其基站

“伪基站”（pseudo base station）即假基站，设备一般由主机和笔记本电脑组成。伪基站设备运行时，用户手机被强制连接到该设备上，导致手机无法正常使用运营商提供的服务，手机用户一般会暂时脱网 8～12s 后恢复正常，部分手机则必须开关机才能重新入网。这个短暂的脱网，就是伪基站实施攻击的机会。这时，伪基站通过短信群发器、短信发信机等

相关设备能够搜取以其为中心、一定半径范围内的手机卡信息，利用 2G 移动通信的缺陷，通过伪装成运营商的基站，冒用他人手机号码强行向用户手机发送诈骗、广告推销等短信息。

**2. 伪基站的主要作案手法**

1）利用 GSM 协议漏洞，冒充其他号码诈骗

GSM 通信协议存在问题，老式的 SIM 卡并没有验证呼叫方是否合法，也无法验证 SMS 短信发送方是否来自真实的手机。伪基站利用这个漏洞，可以随意以某些特别号码，如 10086 或 95588 等发送短信，使手机用户误以为真的是运营商、银行或其他商家发送的短信。以此进行诈骗。《全国首份伪基站短信治理报告》显示，70.2%的诈骗短信冒充运营商诱导用户点击恶意网址；19.4%的诈骗短信冒充银行实施诈骗；4.6%的诈骗短信内容为欺骗用户订低价机票，诱骗回拨电话进行诈骗。报告建议，不要轻信陌生号码发来的短信，更不要随意点击陌生链接；即使是常见的服务号或好友号码，也要对短信内容进行甄别，最好回拨电话进行咨询验证。

2012 年 9 月至 11 月，深圳有一起全国首例入刑的“伪基站”案件，汪某团伙在深圳宝安机场利用一台伪基站设备发送机票广告短信，非法获取 62 万部手机信息，导致机场众多手机用户无法通话，造成区域手机通话业务量损失达 100 余万元。深圳一名司机介绍，他到机场后手机经常信号不好，有一次中断了 2 个小时。

2）注册“钓鱼”网站，实施“钓鱼”式诈骗

钓鱼泛指网络诈骗者利用发送不实信息来诱惑用户上当，从而达到获得用户数据信息的诈骗行为。伪基站“钓鱼”手法与之相似，诈骗者首先会注册大量与运营商或银行等的官网类似的域名，然后将事先注册好的钓鱼域名与钓鱼系统（网站、手机应用）通过伪基站发送。短信的内容一般为“尊敬的用户，因您的话费积分没有兑换即将清零，请登录××网站，下载客户端兑换 287.80 元现金礼包”。然后选择银行卡类型并填写相关信息（姓名、手机、账号、身份证、密码、有效期、CVV2 等），最终这些敏感的信息被入库记录进行洗钱。手机木马负责拦截用户短信，并将如银行交易验证码等关键信息发给远程的黑客。

2015 年 2 月初，市民郑女士在上班时收到 10086 发来的一条短信，称郑女士有大量积分，可以兑换一笔金额不小的话费。“10086 经常发送一些话费信息，而且我看到短信的发送号码是 10086，因此深信不疑。”郑女士表示，她随后点击了短信上附带的网址链接，进入了一个兑换话费的网页，并按提示输入了自己的支付宝账号密码和银行卡密码。郑女士等了几天，说好的话费却迟迟没有到账。更蹊跷的是，她发现自己在支付宝上绑定的三张银行卡内，资金缩水了。经查询资金流水，她发现 2.7 万元被盗。

2015 年 3 月 27 日，国内互联网安全漏洞平台乌云网发布一则名为《一场钓鱼引发的大量网银密码泄露（各大银行均有中招）》的告示说，攻击者利用 10086 伪基站进行钓鱼，通过让用户点击兑换积分实施诈骗。登录该页面可获得用户登录网银、身份证、密码等信息。乌云网“白帽子”破解该网站的后台系统，登录后发现每一个伪基站都掌握 7000 条以上的银行卡数据。根据用户账号可成功登录该用户的网银界面。根据乌云网统计，从伪基站泄露银行卡信息的用户粗略估算有数百万个，2015 年每年大约有十万用户被骗信息。

3）不断变化位置，进行反侦察

伪基站具有很强的流动性，通常将“伪基站”设备放置在汽车内，驾车缓慢行驶或将车停在特定区域，进行短信诈骗或广告推销。例如有一个团伙，就经常在银行附近发送诈骗短信。他们还使用游击战，一般一个地方只待一天，有的团伙甚至转战20多个省市。

4）其他

伪基站还会导致手机用户频繁地更新位置，使得该区域的无线网络资源紧张并出现网络拥塞现象，影响用户的正常通信。

## 1.5 数据驱动型攻击

数据驱动型攻击是通过向某个程序发送数据，以产生非预期结果的攻击，通常为攻击者给出访问目标系统的权限。它分为缓冲区溢出攻击、格式化字符串攻击、整数溢出攻击、悬浮指针攻击、同步漏洞攻击和信任漏洞攻击等。本节介绍出现较多的前面两种。

### 1.5.1 缓冲区溢出攻击

1988年，臭名昭著的Robert Morris蠕虫事件曾造成全世界6000多台网络服务器瘫痪，给这些用户总共带来约200万～6000万美元的损失。从此，Hacker一词开始被赋予了特定的含义，罗伯特·莫里斯的名字也广为人知。他使用的就是缓冲区溢出攻击（buffer overflow attack）。

#### 1. 缓冲区溢出

缓冲区是程序运行时在内存中为保存给定类型的数据而开辟的一个连续空间。这个空间是有限的。当程序运行过程中要放入缓冲区的数据太多时，就会产生缓冲区溢出。请看下面的例子。

**例1.1** 一个C函数。

```
void function(char *str) {
    char buffer[16];
    strcpy(buffer,str);
}
```

为了测试这个函数，可以使用等价分类法，设计两种字符串。

（1）str的长度小于16。

（2）str的长度大于16。

显然，使用测试数据（1），应该没有什么问题；使用测试数据（2），就会造成buffer的溢出，出现Segmentation fault（分段错误）。因为函数strcpy()没有进行变量边界的检查，导致缓冲区溢出了。除了strcpy()外，存在类似问题的标准函数还有strcat()、sprintf()、vsprintf()、gets()和scanf()等。

但是，这时并没有形成攻击，只能算作程序设计不严谨，形成了应用程序的漏洞。

### 2. 缓冲区溢出攻击

利用缓冲区溢出漏洞，黑客可以精心策划两种攻击。

（1）利用缓冲区溢出，关闭某程序或使其无法执行。图 1.44 表明了函数栈帧的结构和在栈中的位置。其中，栈顶指针（Stack Pointer，SP）用于指示栈顶的偏移地址，基数指针（Base Pointer，BP）用于确定在堆栈中的操作数地址。函数调用发生时，新的堆栈帧被压入堆栈；当函数返回时，相应的堆栈帧从堆栈中弹出。对于一个输入来说，如果发生了缓冲区溢出，有可能使过多的输入数据进入内存的其他区域，而这个区域内存放着另外一个程序的代码，这些数据会覆盖这个程序的部分代码，使该程序不能正常运行、错误地关闭或无法执行。这种情形可能发生在本地（称为本地溢出），也可能发生在远程（称为远程溢出）。例如，一个用户要登录远程的某服务器，首先要进行登录，输入密码。如果该服务器的登录程序有缓冲器溢出漏洞，则可能会导致服务器的某些程序关闭。特别是扰乱具有某些特权运行的程序的功能，就可以使攻击者取得程序的控制权。

（2）启动一个恶意代码。如图 1.45 所示，利用缓冲区溢出还有可能使得一个函数返回一个恶意代码的地址。

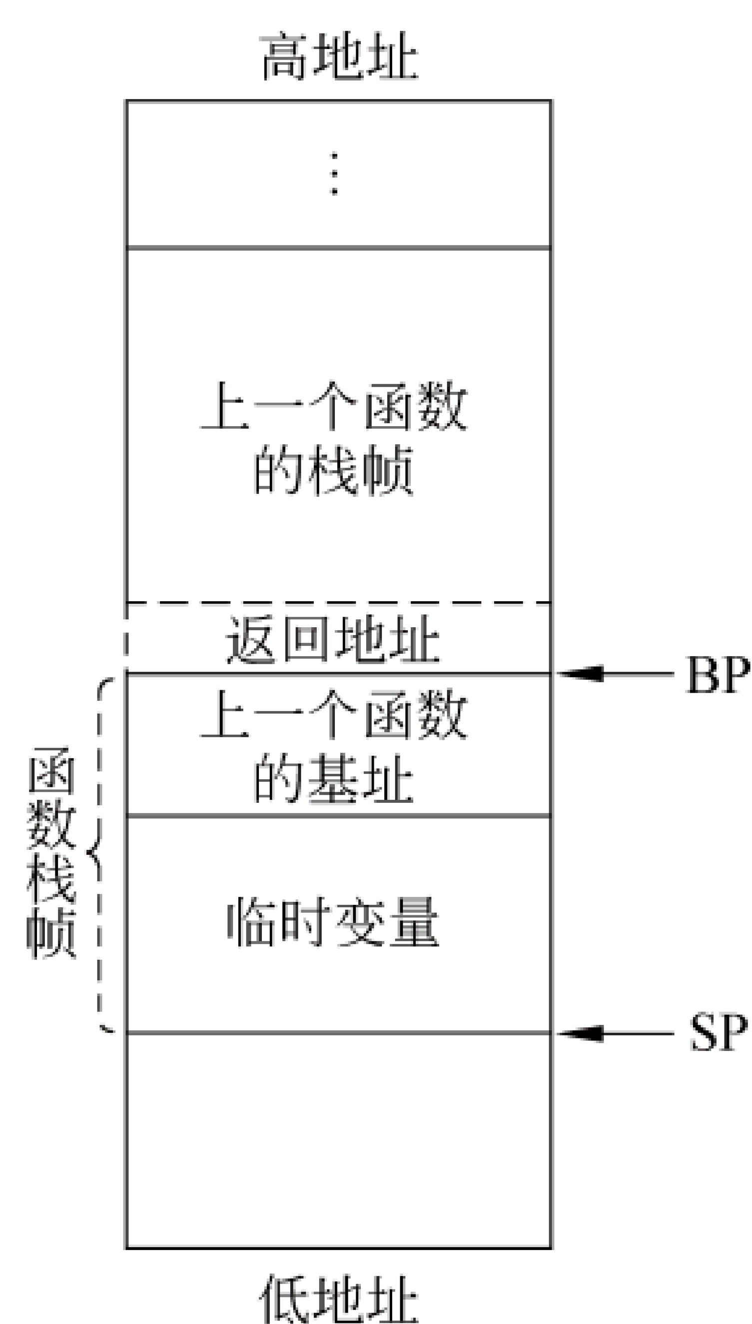



图 1.44　函数栈帧的结构

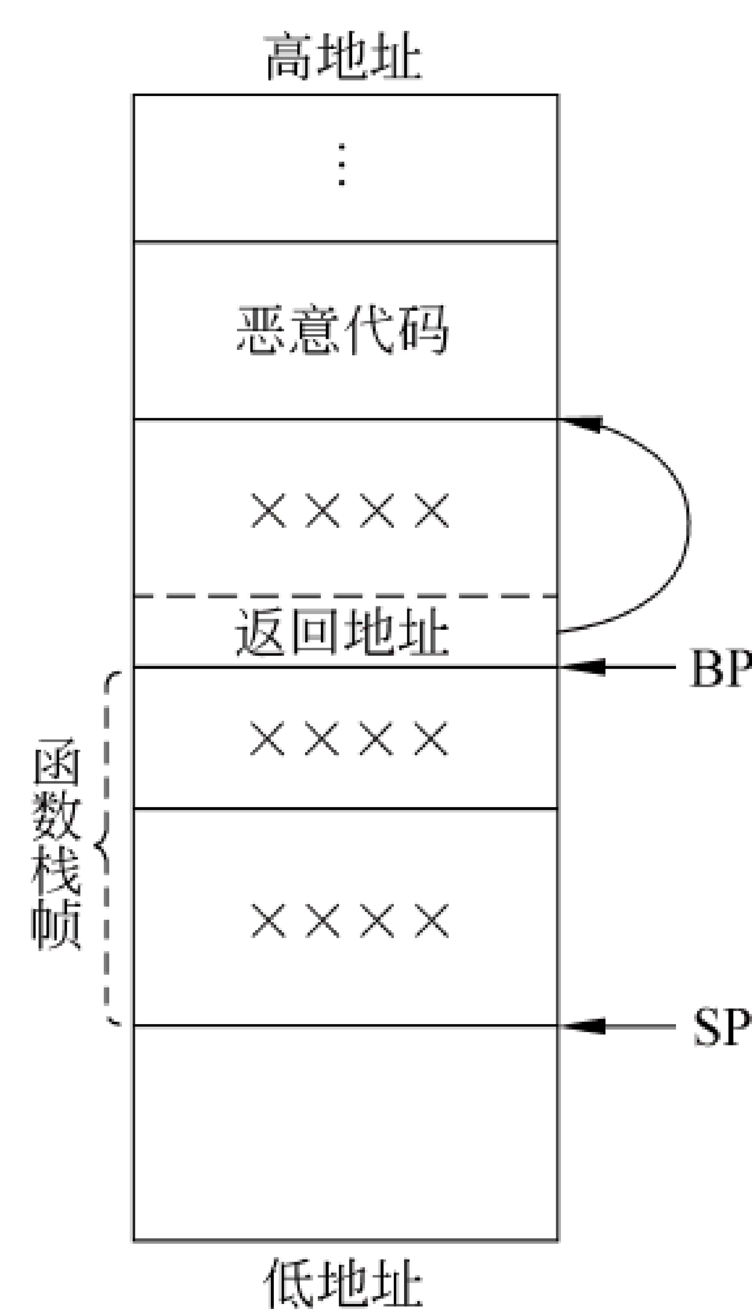



图 1.45　缓冲区溢出攻击

### 3. 缓冲区溢出防御措施

（1）编写安全的代码。尽可能设计和编写安全没有漏洞的程序，避免程序中有不检查变量、缓冲区大小及边界等情况存在。例如，使用 grep 工具搜索源代码中容易产生漏洞的库调用，检测变量的大小和数组的边界，对指针变量进行保护，以及使用具有边界和大小检测功能的程序设计语言编译器等。

（2）基于一定的安全策略设置系统。如攻击者攻击某个 Linux 系统，必须事先通过某些途径对要攻击的系统做必要的了解，如版本信息等，然后再利用系统的某些设置直接或间

接地获取控制权。因此，防范缓冲区溢出攻击就要对系统设置实施有效的安全策略。

减少以 root 权限运行的代码。减少使用 SUID 的 root 程序。这样即使攻击者成功地执行了缓冲区溢出攻击，他们还得继续把自己的特权升级到 root。

（3）保护堆栈。主要有以下两种措施。

① 加入函数建立和销毁代码。前者在函数返回地址后增加一些附加字节，返回时要检查这些字节有无被改动。

② 使堆栈不可执行——非执行缓冲区技术，使入侵者无法利用缓冲区溢出漏洞。

（4）测试并审核每个程序。

（5）安装由厂家提供的所有相关的安全补丁。

### 1.5.2 格式化字符串攻击

格式化字符串攻击与普通缓冲区溢出攻击有一些相似之处，但又有不同。普通的缓冲区溢出攻击利用的是堆栈生长方向与数据存储方向相反，用后存入的数据来覆盖先前压栈的返回地址，从而改变程序预定的流程。而格式化字符串攻击是利用程序中的一些本应指定用户输入格式，却没有严格指定用户输入格式的函数，通过提交特殊的格式字符串进行攻击。

#### 1. 格式化字符串函数族

ANSI C 定义了一系列的格式化字符串函数。

printf：输出到一个 stdout 流。

fprintf：输出到一个文件流。

sprintf：输出到一个字符串。

snprintf：输出到一个字符串并检查长度。

vprintf：从 va_arg 结构体输出到一个 stdout 流。

vfprintf：从 va_arg 结构体输出到一个文件流。

vsprintf：从 va_arg 结构体输出到一个字符串。

vsnprintf：从 va_arg 结构体输出到一个字符串并检查长度。

另外，还有基于这些函数的复杂函数和非标准函数，包括 setproctitle、syslog、err*、verr*、warm 和*vwarm 等。

这些函数有一个共同的特点，即都要求使用一个格式化字符串。例如，对于 printf 函数，它的第一个参数就是格式化字符串。

#### 2. 格式化字符串漏洞

为了说明对格式化字符串使用不当而产生的格式化字符串漏洞，请先看下面的程序。

**例 1.2**

```
#include <stdio.h>
int main() {
   char *name;
```

```
    gets(s);
    printf(s);
}
```

下面是该函数的两次运行结果：

```
abcde
abcde%08x,%08x,%08x
000002e2,0000ffe4,0000011d
```

也就是说，当输入 abcde 时，输出仍然是 abcde。而当输入“%08x,%08x,%08x”时，输出的却是“000002e2,0000fe4,0000011d”。这就是格式化字符串漏洞所造成的问题。因为，在 printf 函数中，s 被解释成了格式化字符串。当调用该函数时，首先会解析格式化字符串，一次取一个字符进行分析：如果字符不是%，就将其原样输出；若字符是%，则其后面的字符就要按照格式化参数进行解析。当输入 abcde 时，由于没有包含%，所以每个字符都被原样输出了。而当输入“%08x,%08x,%08x”时，就要将每个%后面的 x 都解释为一个十六进制的数据项，但函数没有这样 3 个数据项。于是，就将堆栈中从当前堆栈指针向堆栈底部方向的 3 个地址的内容按十六进制输出出来，这就是“000002e2,0000fe4,0000011d”。

这就给人们一个启发：当格式化字符串中包含有许多%时，就会有机会访问到一个非法地址。

### 3. 格式化字符串攻击的几种形式

（1）查看内存堆栈指针开始的一些地址的内容。使用类似于

```
printf ("%08x,%08x,%08x");
```

的语句，可以输出当前堆栈指针向栈底方向的一些地址的内容，甚至可以是超过栈底之外的内存地址的内容。

（2）查看内存任何地址的内容。所查看的内存地址内容也可以从任何一个地址开始的内存内容。例如，语句

```
printf ("\x20\02\x85\x08_%08x,%08x,%08x");
```

将会从地址 0x08850220 开始，查看连续 3 个地址的内容。

（3）修改内存任何地址的内容。格式化字符串函数还可以使用一个格式字符%n。它的作用是将已经打印的字节数写入一个变量。请观察下面的程序。

**例 1.3**

```
#include <stdio.h>
int main() {
    int i;
    printf("china \\%n\\n",(int*)&i);
    printf("i = %d\\n",i);
}
```

程序运行的结果如下：

```
china
i = 5
-
```

即 i 的值为前面已经打印的字符串 china 的长度 5。利用这一点，很容易改变某个内存变量的值。

**例 1.4**

```
#include <stdio.h>
int main(){
   int i = 5;
   printf("%108u%n\\t",1,(int*)&i);printf("i=%d\\n",i);
   printf("%58s123%n\\t","",&i);print("i=%d\\n",i);
}
```

程序执行结果如下：

```
       1    i=108
   123      i=26
```

语句

```
printf("%108u%n\t",1,(int*)&i);
```

用数据 1 的宽度 108 来修改变量 i 的值。而语句

```
printf("%58s123%n\\t"," ",&i);
```

是用字符串" "加上字符串 123 的存放宽度 23+3 来修改变量 i 的值。

使用同样的办法，可以向进程空间中的任意地方写一个字节，以达到下面的目的。

① 通过修改关键内存地址内容，实现对程序流程的控制。

② 覆盖一个程序储存的 UID 值，以降低和提升特权。

③ 覆盖一个执行命令。

④ 覆盖一个返回地址，将其重定向到包含 shell code 的缓冲区中。

# 1.6 拒绝服务攻击

## 1.6.1 拒绝服务攻击及其基本方法

拒绝服务（Denial of Service，DoS）攻击并不是某一种具体的攻击方式，而是攻击所表现出来的结果，其最终使得目标系统因遭受某种程度的破坏而不能继续提供正常的服务，甚至导致物理上的瘫痪或崩溃。具体的攻击方法可以是多种多样的，可以是单一的手段，也可以是多种方式的组合利用，最基本的 DoS 攻击就是利用合理的服务请求来占用过多的服务资源。下面介绍几种典型的拒绝服务攻击。

### 1. IP 碎片攻击

1）IP 碎片与 IP 碎片漏洞

数据链路层对于所传输的帧有一个长度限制，不允许超过最大传输单元（Maximum Transmission Unit, MTU）。不同网络的 MTU 值不相同，以太网的 MTU 为 1500，IEEE 802.3/802.2 为 1492。这个值可以用 netstat -i 查看。

由于在 IP 层中，数据包要由数据部分加上 IP 头（长度为 20）和传输层分组头（UDP 头长度为 8），所以当要传输 UDP 分组时，数据部分只能有 1500–20–8=1472。数据部分大于这个值，就要进行分片（fragmentation）以满足在以太网中的传输要求。但是，数据被分片后，组成一个 IP 包的各分片都到达目的主机时才进行重组。因此，分片会导致传输效率降低。

在 IP 协议规范中规定了一个 IP 包的最大尺寸，而大多数的包处理程序又假设包的长度超过这个最大尺寸这种情况是不会出现的。因此，包的重组代码所分配的内存区域也最大不超过这个最大尺寸。这样，超大的包一旦出现，包中的额外数据就会被写入其他正常区域。这很容易导致系统进入非稳定状态，是一种典型的缓存溢出（buffer overflow）攻击。

2）死亡之 ping（ping of death）攻击

死亡之 ping 是早期使用的一种简单的 IP 分片攻击。ping 是网际控制消息协议（Internet Control Message Protocol，ICMP）中的一个应用程序。ICMP 是一种差错报告机制，可以让路由器或目标主机将遇到的差错报告给源主机，以弥补 IP 协议无连接、无差错报告和差错纠正机制的不足。如图 1.46 所示，ICMP 报文始终包含 IP 首部和产生 ICMP 差错报文的 IP 数据报的前 8 个字节（64KB）。由于这一特点，早期的许多操作系统在处理 ICMP（如接收 ICMP 数据报文）时，只开辟 64KB 的缓存区。在这种情况下，一旦处理的数据报的实际长度超过 64KB，操作系统将会产生一个缓冲溢出，引起内存分配错误，最终导致 TCP/IP 协议堆栈的崩溃，造成主机死机。

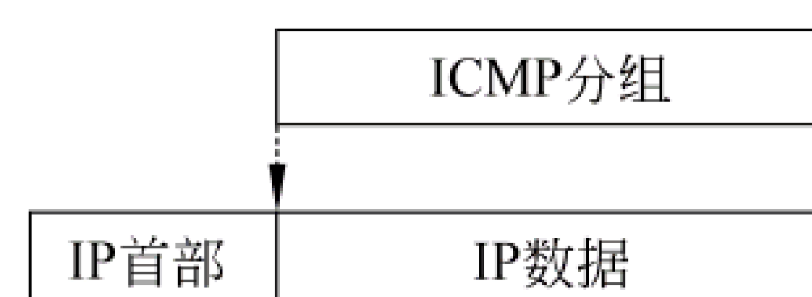


图 1.46 ICMP 分组的封装

ping 通过发送 ICMP 测试包来测试一台主机的可达性。死亡之 ping 进行攻击时，可以执行以下命令：

```
ping -l 65535 目标 IP -t
```

其中：

参数 l（L）用于指定包的长度，这里是 65 535。因为 IP 包头中用于指定 IP 数据包长度的字段为 2B，所以一个 IP 数据包的最大长度为 $2^{16}$=65 536。取 65 535 已是相当大了。

参数 t（T）要求一直 ping。这时，对方的主机若存在这种漏洞，就会形成一次拒绝服务攻击。但是，现在的操作系统所附带的 ping 程序都限制了发送数据包的大小。因而这样的攻击已经不再可能。

**2．“泪滴”（teardrop）**

“泪滴”攻击就是入侵者伪造数据报文，向目标机发送含有重叠偏移的畸形数据分段：第一个包的偏移量为 0，长度为 *N*；第二个包的偏移量小于 *N*……如图 1.47 所示。这样的畸形分片传送到目的主机后，在堆栈中重组时，需要超乎寻常的巨大资源，从而造成系统资源的缺乏，协议栈崩溃。

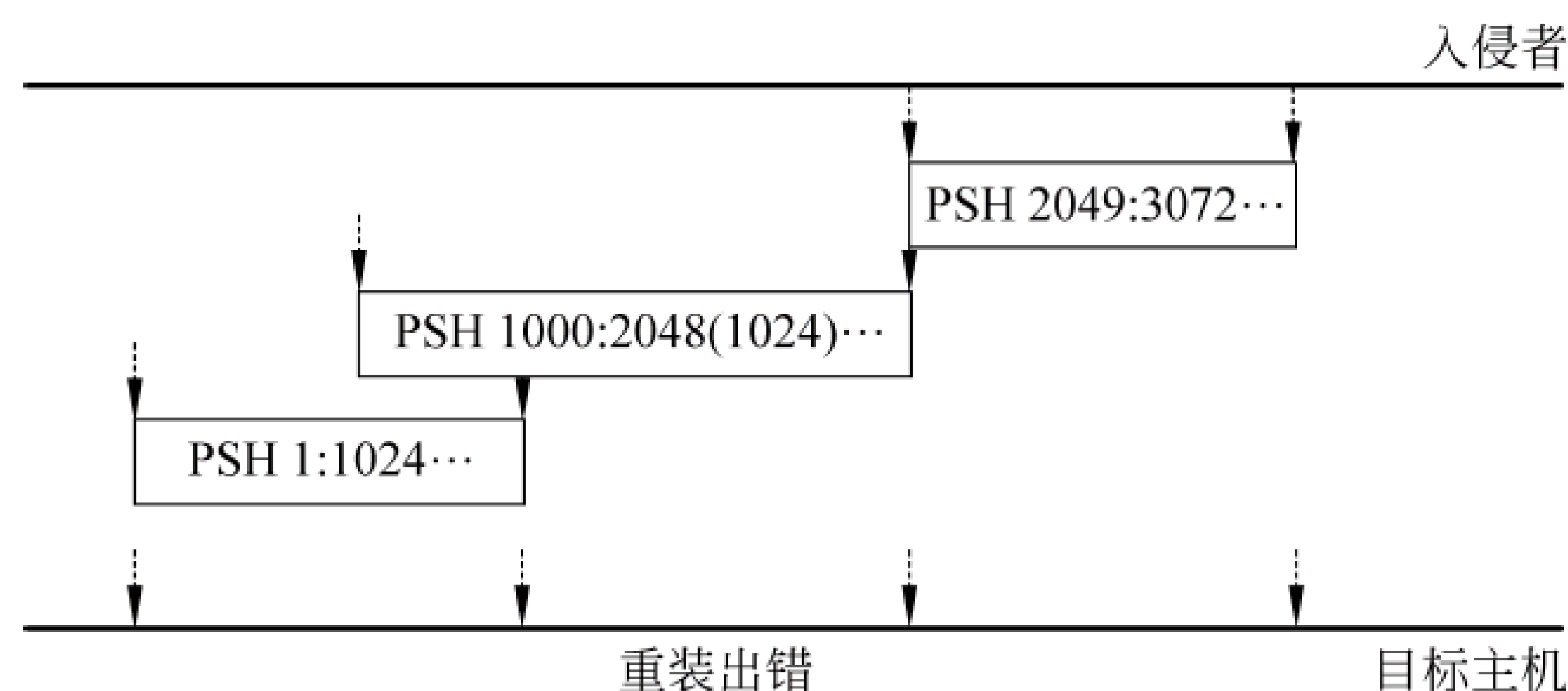



图 1.47　含有重叠偏移的畸形数据分片

**3．UDP“洪水”（UDP flood）**

UDP“洪水”，也称为 UDP 淹没，是基于主机的拒绝服务攻击的一种。其原理非常简单，因为 UDP 是一种无连接的协议，不需要用任何程序建立连接就可以传送数据。这样，攻击者只要开启一个端口提供相关的服务，就可以对攻击对象实施针对相关服务的攻击。常见的情况是利用大量 UDP 小包对 DNS 服务器、Radius 认证服务器、流媒体服务器以及防火墙等发起攻击，造成网络瘫痪。例如，攻击可以针对 Echo/Chargen 服务进行。Echo/Chargen 服务是 TCP/IP 为 UDP 提供的两种服务。Echo 的作用就是由接收端将接收到的数据内容返回到发送端，Chargen 则随机返回字符。这样简单的功能，为网络管理员提供了进行可达性测试、协议软件测试和选路识别的重要工具，也为黑客进行“洪水”攻击提供了方便。当入侵者假冒一台主机向另一台主机的服务端口发送数据时，Echo 服务或 Chargen 服务就会自动回复。两台主机之间的互相回送会形成大量数据包。当多台主机之间相互产生回送数据包时，最终会导致系统瘫痪。

**4．SYN“洪水”（SYN flood）与 Land**

SYN flood 是当前最流行的拒绝服务攻击方式之一。它是一种利用 TCP 协议缺陷，发送大量伪造的 TCP 连接请求，从而使得被攻击方资源耗尽（CPU 满负荷或内存不足）的攻击方式。

这种攻击的基本原理还是要从 TCP 连接建立的过程——三次握手（three-way handshake）说起。这个三次握手过程存在着漏洞：假设一个客户向服务器发送了 SYN 报文后突然死机或掉线，那么服务器在发出 SYN+ACK 应答报文后就无法收到客户端的 ACK 报文，使第三次握手无法完成。而服务器并不知道客户端发生了什么情况，于是就会重试，再次发送 SYN+ACK 给客户端，并等待一段时间——SYN Timeout（大约为 30 秒至 2 分钟）后丢弃这个半连接。这个情况似乎很正常，但只能说在正常情况下很正常。而在非正常情况下，

就会出现问题，因为它使攻击者有机可乘：假如攻击者大量模拟这种情况，服务器端将要维护一个非常大的半连接列表，即便是简单的保存并遍历也会消耗非常多的 CPU 时间和内存，何况还要不断地对这个列表中的 IP 进行 SYN+ACK 的重试。这种情况下，若服务器的 TCP/IP 栈不够强大，最后就会导致堆栈溢出使系统崩溃；即使服务器端的系统足够强大，服务器端也会因忙于处理攻击者伪造的 TCP 连接请求而无暇理睬客户的正常请求，使服务器无法再服务。

Land 也是利用三次握手的缺陷进行攻击；但它不是依靠伪造的地址，而是先发出一个特殊的 SYN 数据包，包中的源地址和目标地址都是目标主机。这样，就会让目标主机向自己回以 SYN+ACK 包，导致自己又给自己回一个 ACK 并建立自己与自己的连接。大量这样的无效连接达到一定数量，将会拒绝新的连接请求。

#### 5．MAC Flood 攻击

MAC Flood 攻击是针对交换机的攻击。在交换式局域网中，利用交换地址映射表，将从一个端口（MAC）接收到的数据转发到另外的端口（MAC）。交换机型号不同，MAC 地址表中可容纳的 MAC 地址数量也不同。在正常情况下，MAC 地址表的容量是足够使用的。另一方面，为了使交换机地址映射表不被过期的地址挤满，交换机常使用动态交换地址映射表，其特点是规定了一个 age time——MAC 地址老化时间，默认为 5 分钟。如果在 age time 没有收到过任何 MAC 地址表条目的数据帧，则将该 MAC 地址条目删除。

利用 MAC 地址映射表进行攻击时，攻击者可以使用一个程序，伪造大量包含随机源 MAC 地址的数据帧发往交换机。由于有些攻击程序 1 分钟就可以发出十几万个伪造的 MAC 地址，而交换机一般 MAC 地址表只有几千条，所以瞬间就会把交换机的 MAC 地址表填满。于是，交换机再接到数据，不管是单播、广播还是组播，都找不到需要的地址表条目，无法找到对应的端口进行转发，只能向所有端口广播。

### 1.6.2 分布式拒绝服务攻击

分布式拒绝服务（Distributed Denial of Service，DDoS）攻击指借助于客户/服务器技术，将多个计算机联合起来作为攻击平台，对一个或多个目标发动 DoS 攻击，从而成倍地提高拒绝服务攻击的威力。通常，攻击者使用一个偷窃账号将 DDoS 主控程序安装在一个计算机上，在一个设定的时间主控程序将与大量代理程序通信，代理程序已经被安装在 Internet 上的许多计算机上。代理程序收到指令时就发动攻击。利用客户/服务器技术，主控程序能在几秒钟内激活成百上千次代理程序的运行。

#### 1．DDoS 系统的一般结构

如图 1.48 所示，一个比较完善的 DDoS 攻击体系分成如下 4 个部分。

（1）攻击者：整个攻击过程的发起者，其所用主机称为攻击主控台，可以是网络上任何一台主机，用来向主控端发送命令。

（2）主控端：攻击者非法侵入并控制的一些主机，其上安装了特殊程序用来接收攻击者的命令，并向它们控制的各代理端发出这些命令。

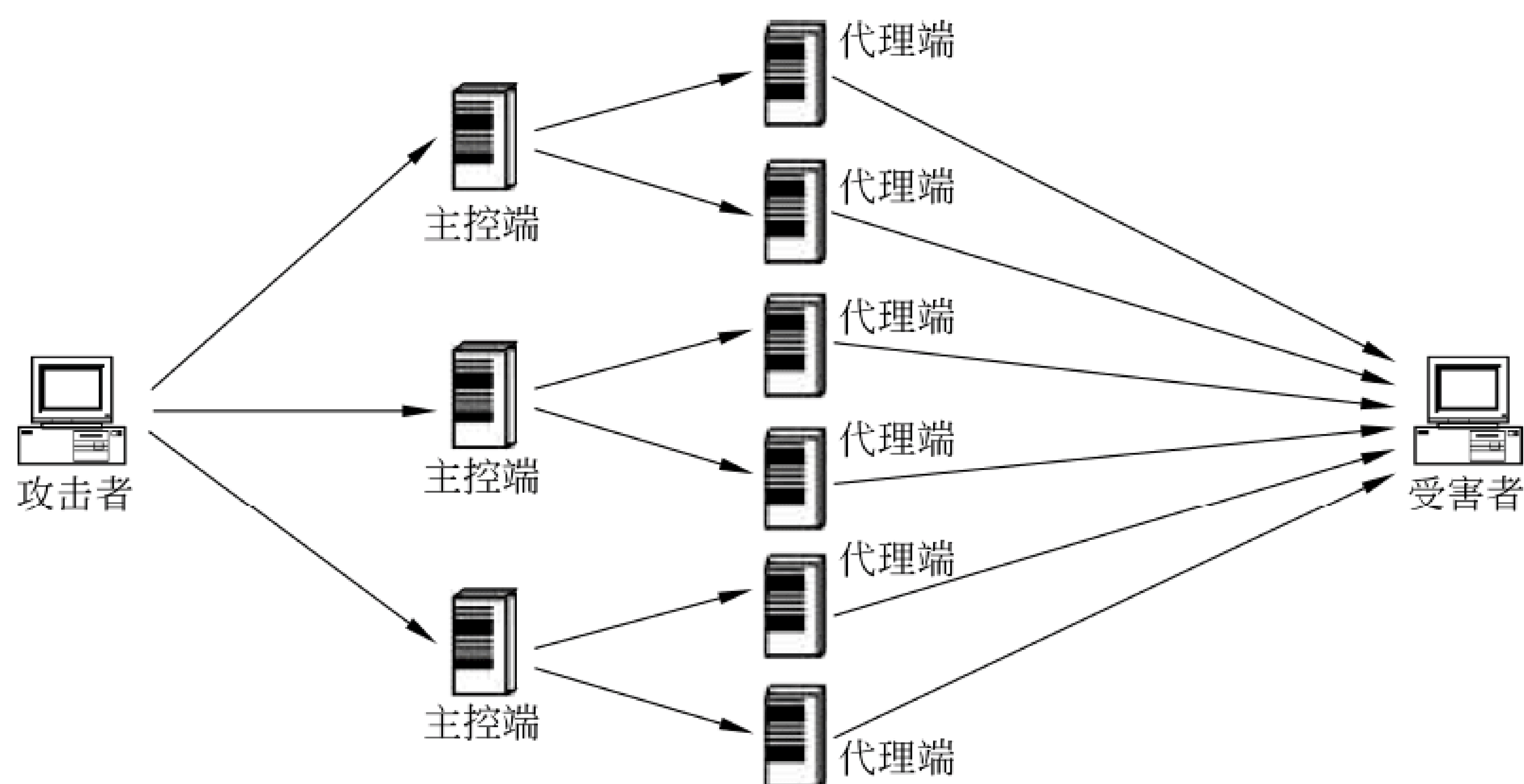


图 1.48　DDoS 攻击的原理

（3）代理端：即傀儡机，也是攻击者控制的一些主机，其上运行攻击程序。

（4）受害者。

**2. 组织一次 DDoS 攻击的过程**

这里用“组织”这个词，是因为 DDoS 并不像入侵一台主机那样简单。一般来说，黑客进行 DDoS 攻击时会经过如下 3 个阶段。

1）搜集了解目标的情况

下列情况是黑客非常关心的情报。

（1）被攻击目标主机数目、地址情况。

（2）目标主机的配置、性能。

（3）目标的带宽。

对于 DDoS 攻击者来说，攻击互联网上的某个站点，有一个重点就是确定到底有多少台主机在支持这个站点，一个大的网站可能有很多台主机利用负载均衡技术提供同一个网站的 WWW 服务。以 Yahoo！为例，一般会有下列地址提供 WWW 服务：

66.218.71.87

66.218.71.88

66.218.71.89

66.218.71.80

66.218.71.81

66.218.71.83

66.218.71.84

66.218.71.86

对一个网站实施 DDoS 攻击，就要让这个网站中所有 IP 地址的机器都瘫痪。因此事先搜集情报对 DDoS 攻击者来说是非常重要的，这关系到使用多少台傀儡机才能达到效果的问题。

2）占领傀儡机

黑客最感兴趣的是有下列情况的主机。

（1）链路状态好的主机。

（2）性能好的主机。

（3）安全管理水平差的主机。

首先，黑客做的工作一般是扫描，随机地或者是有针对性地利用扫描器去发现网络上那些有漏洞的主机，像程序的溢出漏洞、CGI、Unicode、FTP、数据库漏洞等，都是黑客希望看到的扫描结果。随后就是尝试入侵了。

黑客在占领了一台傀儡机后，除了要进行留后门、擦脚印这些基本工作之外，还要把 DDoS 攻击用的程序上传过去，一般是利用 FTP。在攻击机上，会有一个 DDoS 的发包程序，黑客就是利用它来向受害目标发送恶意攻击包的。

3）实际攻击

如果前面的准备做得好，实际攻击过程反而是比较简单的。这时候埋伏在攻击机中的 DDoS 攻击程序就会响应主控台的命令，一起向受害主机以高速度发送大量的数据包，导致它死机或无法响应正常的请求。黑客一般会以远远超出受害方处理能力的速度进行攻击。高明的攻击者还要一边攻击一边用各种手段来监视攻击的效果，以便需要的时候进行一些调整。较为简单的办法就是开一个窗口不断地 ping 目标主机，在能接到回应的时候就再加大一些流量或者命令更多的傀儡机加入攻击。

### 3．DDoS 的监测

现在网上 DDoS 攻击日益增多，只有及时检测，及早发现自己受到攻击，才能避免遭受惨重的损失。检测 DDoS 攻击的主要方法有以下几种。

1）根据异常情况分析

异常情况包括：

（1）网络的通信量突然急剧增长，超过平常的极限值时；

（2）网站的某一特定服务总是失败；

（3）发现有特大型的 ICP 和 UDP 数据包通过，或者数据包内容可疑。

2）使用 DDoS 检测工具

扫描系统漏洞是攻击者最常进行的攻击准备。目前市面上的一些网络入侵检测系统可以杜绝攻击者的扫描行为。另外，一些扫描器工具可以发现攻击者植入系统的代理程序，并可以把它从系统中删除。

### 4．DDoS 实例

1）Smurf 与 Fraggle

将一个目的地址设置成广播地址（以太网地址为 FF:FF:FF:FF:FF:FF:FF）后，将会被网络中所有主机接收并处理。显然，如果攻击者假冒目标主机的地址发出广播信息，则所有

主机都会向目标主机回复一个应答使目标主机淹没在大量信息中，无法提供新的服务。Smurf 和 Fraggle 攻击就是利用广播地址的这一特点将攻击放大而实施的拒绝服务攻击。其中，Smurf 是用广播地址发送 ICMP ECHO 包，而 Fraggle 是用广播地址发送 UDP 包。

显然，Smurf 为了能工作，必须要找到攻击平台，这个平台就是其路由器上启动了 IP 广播功能的系统，这样就能允许 Smurf 发送一个伪造的 ping 信息包，然后将它传播到整个计算机网络中。因而，为防止系统成为 Smurf 攻击的平台，要禁止所有路由器上 IP 的广播功能（一般来讲，IP 广播功能并不需要）。但是，攻击者若从 LAN 内部发动一个 Smurf 攻击，在这种情况下，禁止路由器上的 IP 广播功能就没有用了。为了避免这样一个攻击，许多操作系统都提供了相应设置，防止计算机对 IP 广播请求做出响应。

挫败一个 Smurf 攻击的最简单方法对边界路由器的回音应答（echo reply）信息包进行过滤，然后丢弃它们，使网络避免被淹没。

2）trinoo

trinoo 是复杂的 DDoS 攻击程序，它使用了主控程序对实际实施攻击的任何数量的代理程序实现自动控制。图 1.49 形象地表明了其攻击原理。图中的“傀儡机”就是一些代理，“控制傀儡机”就是安装有主控程序的计算机。

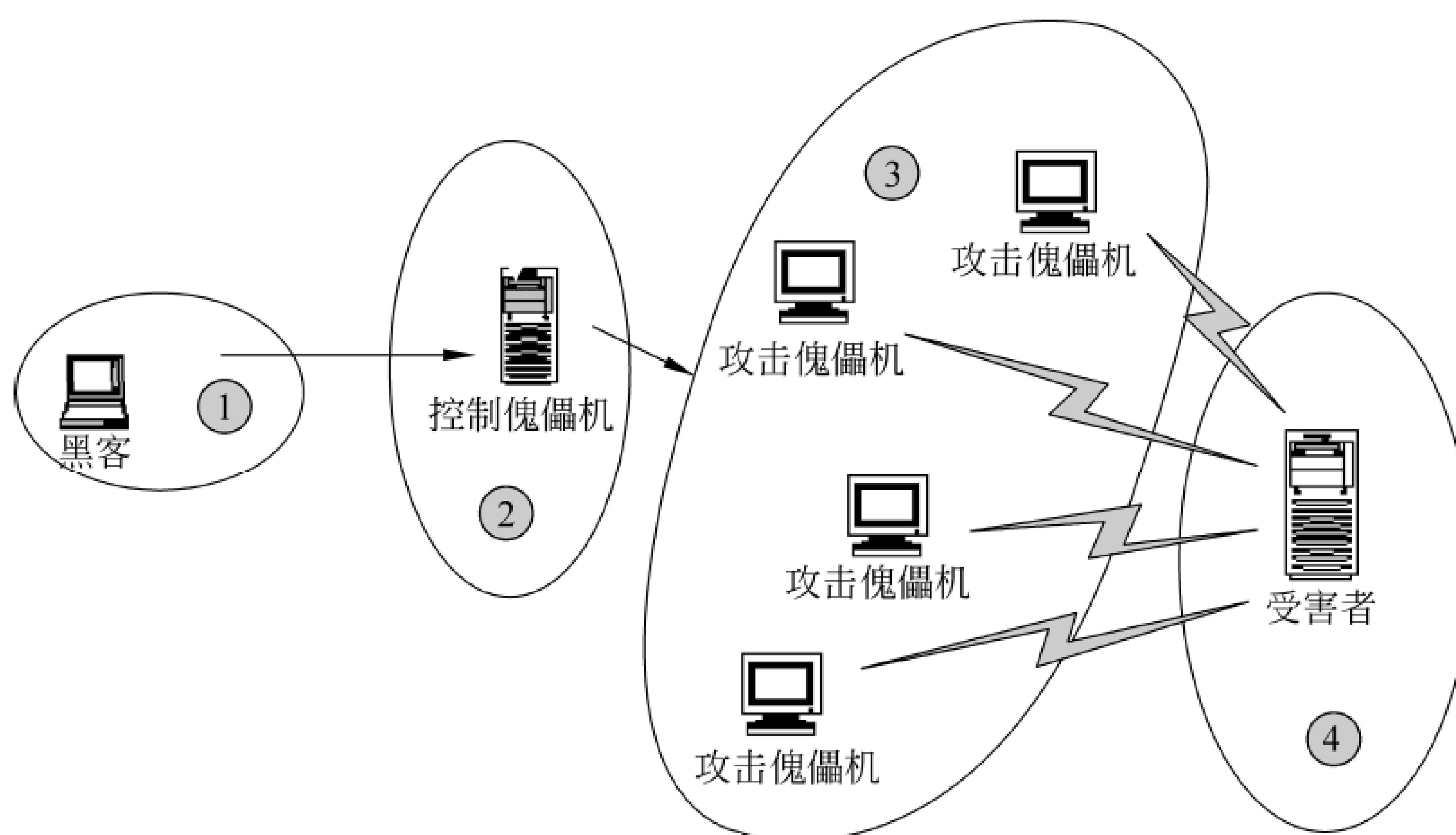



图 1.49　trinoo DDoS 攻击的原理

trinoo DDoS 攻击的基本过程是：攻击者连接到安装了主控程序的计算机，启动主控程序，然后根据一个 IP 地址的列表，由主控程序负责启动所有的代理程序。接着，代理程序用 UDP 信息包冲击网络，攻击目标。在攻击之前，侵入者为了安装软件，已经控制了装有主控程序的计算机和所有装有代理程序的计算机。

DDoS 就是利用更多的傀儡机来发起进攻，以更大的规模来进攻受害者的。

3）Tribe Flood Network 和 TFN2K

Tribe Flood Network 与 trinoo 一样，使用一个主控程序与位于多个网络上的攻击代理进行通信。TFN 可以并行发动数不胜数的 DoS 攻击，类型多种多样（如 UDP 攻击、TCP SYN 攻击、ICMP 回音请求攻击以及 ICMP 广播），而且还可建立带有伪装源 IP 地址的信息包。

TFN2K 是 TFN 的一个更高级的版本，它“修复”了 TFN 的某些缺点。

4）Stacheldraht

Stacheldraht 也是基于 TFN 的，它采用和 trinoo 一样的客户/服务器模式，其中主控程序与潜在的成千个代理程序进行通信。在发动攻击时，侵入者与主控程序进行连接。Stacheldraht 增加了以下新功能：攻击者与主控程序之间的通信是加密的，并使用 rcp（remote copy，远程复制）技术对代理程序进行更新。

### 5. DDoS 攻击的防御策略

DDoS 攻击仍然是我国互联网面临的严重安全威胁之一。其隐蔽性极强，并且攻击的方式和手段不断发生变化。自 2014 年起，利用互联网传输协议的缺陷发起的反射型 DDoS 攻击日趋频繁，增加了攻击防御和溯源的难度。几乎不需要技术基础即可使用的 DDoS 攻击服务平台在互联网上大量出现，DDoS 攻击以服务形式在互联网上公开叫卖，这些平台的出现极大地降低了 DDoS 攻击技术门槛，使攻击者可以轻易发起大流量攻击。2015 年前三季度，攻击流量在 1Gbps 以上的 DDoS 攻击次数近 38 万次，日均攻击次数达到 1491 次。所以加强安全防范意识、提高网络系统的安全性还是当前最为有效的办法。可采取的安全防御措施有以下几种。

（1）及早发现系统存在的攻击漏洞，及时安装系统补丁程序。对一些重要的信息（例如系统配置信息）建立和完善备份机制。对一些特权账号（例如管理员账号）的密码设置要谨慎。通过这样一系列的举措可以把攻击者的可乘之机降低到最小。

（2）在网络管理方面，要经常检查系统的物理环境，禁止那些不必要的网络服务。建立边界安全界限，确保输出的包受到正确限制。经常检测系统配置信息，并注意查看每天的安全日志。

（3）利用网络安全设备（如防火墙）来加固网络的安全性，配置好它们的安全规则，过滤所有可能的伪造数据包。

（4）与网络服务提供商协调工作，请其帮助实现路由的访问控制和对带宽总量的限制。

（5）当发现自己正在遭受 DDoS 攻击时，应当立即启动应急策略，尽可能快地追踪攻击包，并且要及时联系 ISP 和有关应急组织，分析受影响的系统，确定涉及的其他节点，从而阻挡来自已知攻击节点的流量。

（6）发现自己的计算机被攻击者用作主控端和代理端时，不能因为自己的系统暂时没有受到损害而掉以轻心，因为攻击者已发现系统的漏洞，这是一个很大的潜在威胁。同时，一旦发现系统中存在 DDoS 攻击的工具软件要及时把它清除，以免留下后患。

## 实验 4　拒绝服务攻击演示

### 1. 实验目的

了解拒绝服务攻击的种类及其攻击要点。

### 2. 实验内容

（1）总结当前已经发现的各种拒绝服务攻击及其攻击原理（要点）。
（2）针对一种可能造成拒绝攻击的系统漏洞，为其设计一个测试程序。

### 3. 实验准备

（1）收集当前已经发现的各种拒绝服务攻击及其攻击原理（要点）。
（2）收集拒绝服务攻击工具，分析其攻击的机理和利用的系统漏洞。
（3）根据分析的拒绝服务工具所利用的系统漏洞，为其设计一个测试程序。
（4）写出用自己设计的程序和工具进行上述拒绝服务攻击实验的环境及步骤。
（5）制定实验应急预案。

### 4. 实验范例

Linux 的 UNIX 域名套接字没有考虑/proc/sys/net/core/wmem_max 参数的限制，本地普通用户可以通过向某个套接字传送大量数据，导致 Linux 内核分配内存空间时出错，系统停止响应。此时必须重新启动系统。

对该漏洞，可以使用下面的测试程序：

```
#include <sys/types.h>
#include <sys/socket.h>
#include <string.h>

char buf[128 * 1024];

int main (int argc, char**argv)
{
    struct sockaddr SyslogAddr;
    int LogFile;
    int bufsize = sizeof(buf)-5;
    int i;

    for ( i = 0; i < bufsize; i++ )
        buf[i] = ' '+(i%95);
    buf[i] = '';

    SyslogAddr.sa_family = AF_UNIX;
    strncpy (SyslogAddr.sa_data, "/dev/log",sizeof(SyslogAddr.sa_data));
    LogFile = socket (AF_UNIX, SOCK_DGRAM, 0);
    sendto (LogFile, buf, bufsize, 0, &SyslogAddr,sizeof(SyslogAddr));
    return 0;
}
```

### 5. 推荐的分析讨论内容

（1）如何防止和发现拒绝服务攻击？

（2）其他发现或想到的问题。

### 1.6.3 僵尸网络

僵尸（zombie）指人死后的尸体在某种作用下重新起立行走，撕咬活人；被咬者遭受传染，不久也会变成僵尸。僵尸网络（botnet）是指采用一种或多种传播手段，使大量主机感染 bot 程序（僵尸程序），从而在控制者和被感染主机之间形成一个可一对多控制的网络。攻击者通过各种途径传播僵尸程序感染互联网上的大量主机，而被感染的主机将通过一个控制信道接收攻击者的指令，组成一个僵尸网络。

#### 1. 僵尸网络的基本功能

1）作为黑客发动 DDoS 攻击的工具

僵尸网络主要被黑客作为发起 DDoS 攻击的傀儡。攻击的目标可以是 Internet 上任何可用的服务器，但以攻击 Web 服务器数量为最多，通过功能滥用的攻击，例如针对电子公告栏运行能耗尽资源的查询或者在受害网站上运行递归 HTTP 洪水攻击。递归 HTTP 洪水指的是僵尸工具从一个给定的 HTTP 链接开始，以递归的方式顺着指定网站上所有的链接访问，这也称为蜘蛛爬行。

僵尸网络也可用于攻击互联网中继聊天（Internet Relay Chat，IRC）网络，流行的攻击方式是“克隆攻击”。在这种攻击中，控制者命令每个僵尸工具连接大量的 IRC 受害终端。被攻击的 IRC 服务器被来自数千个僵尸工具或者数千个频道的请求所淹没。通过这种方式，受到攻击的 IRC 网络可被类似于 DDoS 攻击击垮。

2）发送垃圾邮件

有些僵尸工具会在一台已感染的主机上打开 Socks 代理，然后让这台主机执行很多恶毒任务，例如发送垃圾邮件等。若一个僵尸网络中有上千个僵尸工具，攻击者就可以发送大量垃圾邮件。有些僵尸工具也执行特殊的功能，如收集电子邮件地址、发送钓鱼（phishing）邮件等。

3）信息窃取

僵尸工具也可用数据包监听器来观察通过一台已被攻陷主机上的明文数据，从中提取敏感数据，例如用户名、密码以及其他一些令人感兴趣的数据。

如果被攻陷主机使用加密的通信通道，也可以安装键盘记录器来获取敏感信息。

如果一台主机不止一次被攻陷并属于多个僵尸网络，包监听就有可能偷窃、收集另一个僵尸网络的关键信息。

4）扩散、升级或下载恶意软件

由于所有的僵尸工具都可以通过 HTTP 或者 FTP 下载并执行，因此非常容易被用于扩散新的僵尸工具或电子邮件病毒。一个拥有一万台用于扩散电子邮件病毒的基础主机的僵尸网络可使扩散非常快并且造成极大的危害。

5）伪造点击量，骗取奖金，操控网上投票和游戏，被网络推手作为绑架舆论的工具

僵尸网络也可被用于获取金钱。这可以通过在主机上安装一个有广告的虚假网站，网站的操作员和一些主机公司协商给点击广告付费。在僵尸网络的帮助下，点击可以自动化，让数千僵尸工具点击弹出广告。如果僵尸工具劫持了攻陷主机的起始页面，当受害者使用浏览器的时候点击就被执行。例如滥用 Google 的 AdSense 程序就是一个很典型的骗取奖金的实例。攻击者可以滥用 AdSense 程序，通过让僵尸网络以自动化的方式点击 Google 中的某些广告和人工提高点击数。

在线投票和游戏越来越引起人们的注意，用僵尸网络来操控它们比较简单。由于每个僵尸工具有不同的 IP 地址，每一票与真人投的票有着相同的可信性。

网络推手也可以利用僵尸网络的优势，滥发留言，绑架社会舆论。

6）下载文件

僵尸工具按照黑客的指示，从指定主机中下载各种文件。

7）启动或终止进程

僵尸工具按照黑客的指示，启动或终止指定进程。

**2. 僵尸程序及其传播**

第一个僵尸程序是 1993 年被开发出来的 EggDrop.bot，用于管理员不在时保护聊天频道。1999 年 11 月被木马 SubSeven 2.1 利用，成功地运用 IRC 协议控制感染了 SubSeven 的计算机。僵尸网络就是以此为开端，随着在木马中大量应用而形成的，并把其中被感染了僵尸的计算机称为 zombie。

僵尸程序本身并不具传播性，而要像木马一样被植入。一般说来，僵尸程序可以借助如下几种方式传播。

1）利用系统漏洞

利用系统漏洞是一种主动传播方式。采用这种方式传播时，首先要用某种扫描工具对一定范围内的计算机进行漏洞扫描，然后获得访问权，并在 Shellcode 执行僵尸程序注入代码。这些漏洞多数都是缓存区溢出漏洞。下面以 Slapper 为例，简单描述这种基于 P2P 协议的僵尸程序的传播过程。

（1）感染 Slapper 的主机，用非法的 GET 请求包扫描相邻网段的主机，希望获得主机的指纹（操作系统版本、Web 服务器版本）。

（2）一旦发现有 Apache SSL 缓存溢出漏洞的主机，就开始发动攻击。攻击者首先在建立 SSLv2 连接时，故意设置一个过大的参数，代码没有对参数做边界检查，并复制该参数到一个堆定位的 SSL_SESSION 数据结构中的固定长度缓冲区，造成缓冲区溢出。手工制作的字段是缓存溢出的关键。漏洞探测者小心翼翼地覆盖这些数据域，不会严重影响 SSL 握手。

2）利用邮件、即时消息通信携带

这与传播木马、蠕虫的方法相同。例如，2005 年“性感鸡”（Worm.MSNLoveme）爆

发就是通过 MSN 消息传播的。

3）伪装软件

很多僵尸程序被夹杂在 P2P 共享文件、局域网内共享文件、免费软件、共享软件和恶意网站脚本中，通过伪装，引诱用户下载、打开或点击，进行僵尸程序传播。

4）利用蠕虫携带

将僵尸程序隐藏在蠕虫代码中进行传播。

### 3. 僵尸程序与黑客之间的通信

僵尸程序与黑客之间的通信通常采用两种方法。

（1）利用已有的通信协议，例如 IRC 协议、P2P 协议、AOL（American Online，美国在线）等，直接加以利用或简单改造。

（2）定制一个私有协议，利用公共端口进行掩护，例如利用最常用的 80 端口传递私有协议。

### 4. 僵尸网络的形成

bot 是英文单词 robot（机器人）的缩写，指这类程序可以自动执行预定义的功能，甚至有一定的智能交互能力，可以在特定情况下完成操纵者赋予的特定任务。

僵尸程序一旦被植入，就会自动执行，主动连接到黑客在僵尸代码中指定的计算机。这台计算机可以是黑客自己的计算机，也可以是黑客作为跳板的计算机——这样黑客更为安全。这样，僵尸程序就可以与黑客依靠一定的协议进行通信了。如图 1.50 所示，当黑客用此方法控制了多台计算机时，就形成了一个僵尸网络。

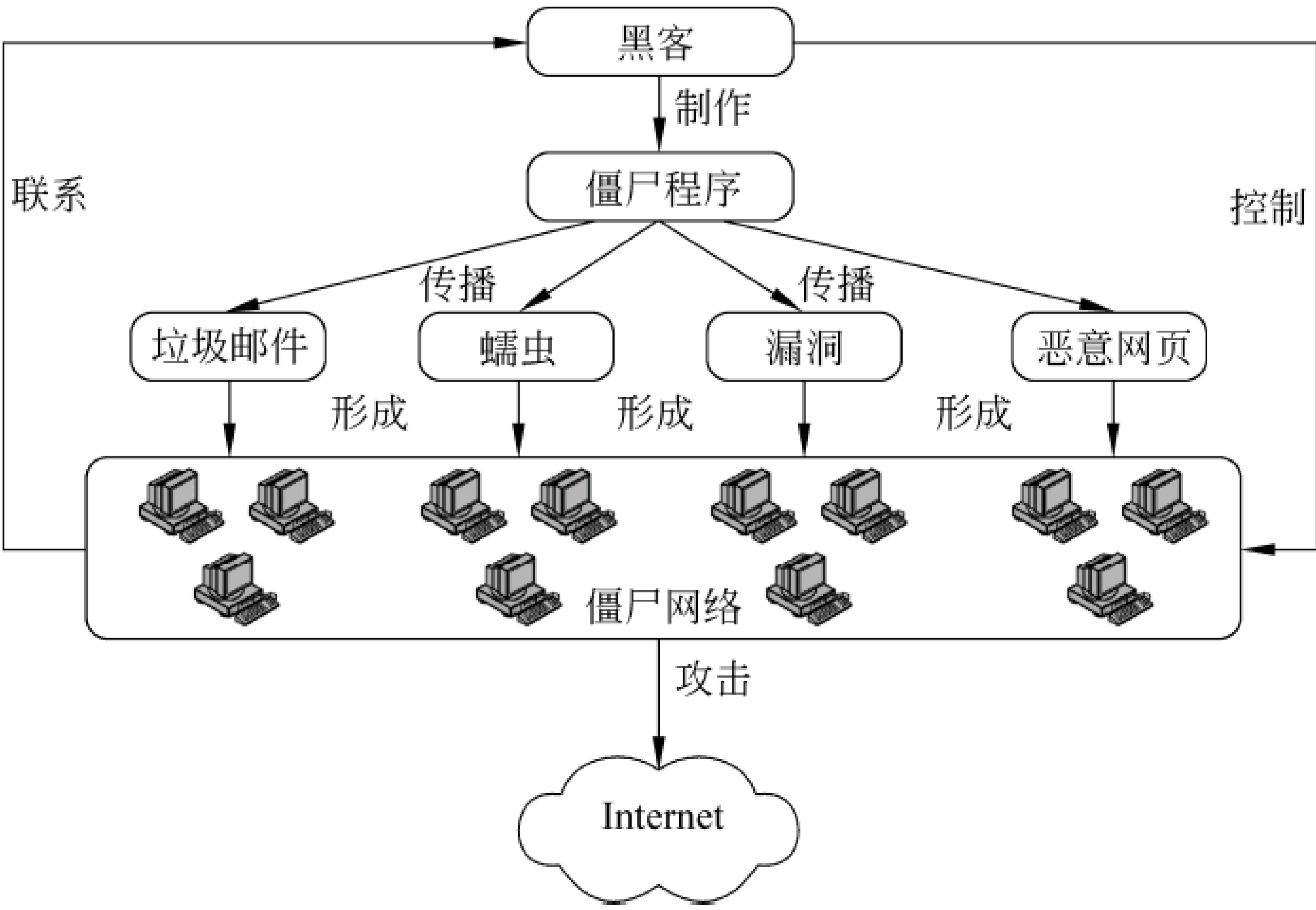



图 1.50　僵尸网络的形成

### 5. 僵尸网络的加入

不同类型的僵尸主机，加入僵尸网络的方式也不同，下面以基于 IRC 协议的僵尸为例，

介绍僵尸主机加入僵尸网络的过程。

（1）如果僵尸中有域名，先解析域名，通常采用动态域名。

（2）僵尸主机与 IRC 服务器建立 TCP 连接。为增强安全性，有的 IRC 服务器设置了连接密码。连接密码在 TCP 三次握手后，通过 PASS 命令发送。

（3）僵尸主机与 IRC 服务器发送 NICK 和 USER 命令，NICK 通常有一个固定的前缀，如 CHN!2345、[Nt]-15120、ph2-1234，前缀通常为国家简称、操作系统版本等。

（4）加入预定义的频道。频道名一般硬编码在僵尸程序体内，为增强安全性，有的控制者为频道设定了密码。中国国家互联网应急中心（CNCERT/CC）的监测数据表明，规模较大（控制 1 万台以上计算机）的僵尸网络通常设置了频道密码，但设置服务器连接密码的僵尸网络还是少数。

### 6. 黑客对于僵尸主机的控制

僵尸网络的主人必须保持对僵尸主机的控制，才能利用它们完成预定的任务目标。下面依然以 IRC 僵尸为例，简单描述一下控制主机对僵尸主机的控制过程。

（1）攻击者或者僵尸网络的主人建立控制主机。大多数控制主机建立在公共的 IRC 服务上，这样做是为了将控制频道做得隐蔽一些；也有少数控制主机是攻击者自己单独建立的。

（2）僵尸主机主动连接 IRC 服务器，加入到某个特定频道。此过程在上面已经介绍了。

（3）控制者（黑客）主机也连接到 IRC 服务器的这个频道上。

（4）控制者（黑客）使用 login、!logon、!auth 等命令认证自己，服务器将该信息转发给频道内所有的僵尸主机，僵尸程序将该密码与硬编码在文件体内的密码比较，相同则将该用户的 nick 名称记录下来，以后可以执行该用户发送的命令。控制者具有 channel op 权限，只有他能发出命令。

### 7. 主控者向僵尸主机发布命令的方法

在 IRC 僵尸网络中，主控者向僵尸主机发送的命令按照要求僵尸程序实现的功能可以分为以下几类。

（1）僵尸网络控制命令。

（2）扩散传播命令。

（3）信息窃取命令。

（4）下载与更新命令。

（5）主机控制命令。可细分为发动 DDoS 攻击、架设服务、发送垃圾邮件和点击欺诈等。

基于 IRC 协议，主控者向受控僵尸程序发布命令的方法有如下 3 种。

（1）设置频道主题（Topic）命令。当僵尸程序登录到频道后，立即接收并执行这条频道主题命令。

（2）使用频道或单个僵尸程序发送 PRIVMSG 消息。这种方法最为常用，即通过 IRC 协议的群聊和私聊方式向频道内所有僵尸程序或指定僵尸程序发布命令。

（3）通过 NOTICE 消息发送命令。这种方法在效果上等同于发送 PRIVMSG 消息，但

在实际情况中并不常见。

# 1.7 信息系统风险与安全策略

## 1.7.1 风险=脆弱性+威胁

系统风险是指系统遭受意外损失的可能性，它主要来自系统可能遭受的各种威胁、系统本身的脆弱性以及系统对于威胁的失策。

### 1. 信息系统脆弱性的表现

信息系统的脆弱性表现为安全漏洞（也称为 bug）。如前所述，计算机的安全漏洞是全方位的，也是动态的。下面介绍几个主要的方面。

1）芯片的脆弱性

安全漏洞不仅存在于软件之中，还存在于硬件之中，特别是芯片中也可能存在漏洞。1997 年，法新社在一篇报道中就引用了 Intel 公司发言人汤姆•沃尔德的一段话："我们已经确认奔腾和具有多媒体扩展（MMX）技术的奔腾处理器芯片存在一处新的缺陷。"这个缺陷导致当操作者取得特权发出一个特殊指令时系统死机。

2）操作系统安全漏洞

操作系统是对计算机系统的软硬件资源进行管理、控制的大型综合软件，是计算机系统运行的基础，操作系统的不安全是计算机系统不安全的重要原因。根据国际权威组织 SANS 和 FBI 于 2003 年公布的安全漏洞报告，在 Internet 的安全漏洞中，排在前 20 名的几乎都是操作系统的漏洞。

从理论上说，任何实际运行的操作系统都会有各自的漏洞。下面列举操作系统的脆弱性的一些共性。

（1）后门式漏洞。后门，或称陷门（trap doors），是一种操作系统的无口令入口，由一段程序实现，通常是系统开发者为调试、测试、维修而设置的简便入口。例如，在特定的时间按下特定的键或提供特定的参数，就会对预定的事件或事件序列产生非授权的影响。发现后门是非常困难的。因此，攻击者也常挖空心思地设计后门，形成隐蔽的信道监视系统运行或伺机对系统发起攻击。例如，操作系统提供的调试器（debug）、向导（wizard）以及 daemon 软件，都有可能被攻击者利用进入系统。

（2）补丁式漏洞。操作系统支持动态连接。因此，操作系统才可以动态地安装 I/O 驱动程序和其他系统服务，也才能通过打补丁的方式修补安全漏洞。当然，这也就为攻击者提供了用打补丁的方式来破坏系统的便利。

（3）远程创建进程式漏洞。操作系统允许远程进程的创建和激活。由于被创建的进程可以继承创建进程的权力，就为攻击者在远程安装攻击软件提供了可能。例如，攻击者可以在远程把补丁打在一个特权用户上，使用这种特权对系统进行攻击。

3）数据库的安全脆弱性

当前，设计数据库系统时主要考虑的内容是数据的共享性、一致性、完整性和访问的可控制性，对于安全的考虑较少。这使数据库系统表现得比较脆弱。例如：

（1）数据库中存放着大量数据。这些数据中重要性、机密性各不相同，而它们却要被不同职责和权力的用户共享，这是十分不安全的。

（2）数据库数据的共享性可能导致一个用户对数据的修改影响了其他用户的正常使用。

（3）数据库一般不保存历史数据，一个数据被修改，旧值就被破坏。

（4）联机数据库可以被多用户共享，可能会造成多个用户操作而破坏数据的完整性。

4）计算机网络的安全脆弱性

计算机网络是通信技术与计算机技术相结合的产物，它的脆弱性主要表现在如下几个方面。

（1）传输中的脆弱性。例如电磁辐射、串音干扰等。

（2）网络体系结构的开放性带来的脆弱性。一个计算机网络要连接多个用户，这本身就是一个不安全因素。特别是对于目前已经普遍使用的 TCP/IP 来说，由于当初主要考虑的是网络互联和传输效率的问题，没有很好地解决安全问题，所以安全的薄弱环节较多。

（3）网络服务的安全脆弱性。例如 Web 服务、FTP 服务、电子邮件服务、DNS 服务、路由服务和 Telnet 服务等，都分别存在自己的漏洞或安全问题。

**2. 信息系统威胁手段**

对于信息系统的威胁有许多方法或手段。下面介绍几种主要的威胁方法。

1）信息窃取

信息窃取的主要途径有以下几个。

（1）在传输中被利用电磁辐射或搭接线路的方式窃取。

（2）授权者向未授权者泄露。例如，一个公司职员在用文件名传输公司秘密文件的同时对文件名编码，使公司的正常秘密文件传输信道被乱用为隐蔽的泄密信道。

（3）存储设备被盗窃或盗用。

（4）未授权者利用特定的工具捕获网络中的数据流量、流向、通行频带和数据长度等数据并进行分析，从中获取敏感信息。

2）扫描（scan）

扫描是利用特定的软件工具向目标发送特制的数据包，对响应进行分析，以了解目标网络或主机的特征。

3）入侵（intrusion）

入侵即非授权访问，是指没有经过授权（同意）就获得系统的访问权限或特权，对系统进行非正常访问，或擅自扩大访问权限越权访问系统信息。主要的非授权访问形式有以下几种。

（1）旁路控制。攻击者利用系统漏洞绕过系统的访问控制而渗入系统内部。

（2）假冒。某个未经授权的实体通过出示伪造的凭证骗取某个系统的信任，非法取得系统访问权或得到额外的特权。

（3）口令破解。利用专门的工具穷举或猜测用户口令。

（4）合法用户的非授权访问。合法用户进入系统后擅自扩大访问权限或越权访问。

4）拒绝服务（DoS）

DoS 指系统可用性因服务中断而遭到破坏。DoS 攻击常常通过用户进程消耗过多的系统资源造成系统阻塞或瘫痪。

5）抵赖（否认）

通信一方由于某种原因而实施的下列行为都称为抵赖。

（1）发方事后否认自己曾经发送过某些消息。

（2）收方事后否认自己曾经收到过某些消息。

（3）发方事后否认自己曾经发送过某些消息的内容。

（4）收方事后否认自己曾经收到过某些消息的内容。

6）滥用（misuse）

滥用泛指一切对信息系统产生不良影响的活动。滥用主要有以下几种。

（1）传播恶意代码。恶意代码是一些对于系统有副作用的代码。它们或者独立存在（如蠕虫）或者依附于其他程序（如病毒、特洛伊木马、逻辑炸弹等），通过大量复制消耗系统资源，或进行删除、修改等破坏性操作，或执行窃取敏感数据的任务。

（2）复制/重放。攻击者为了达到混淆视听、扰乱系统的目的，常常先记录系统中的合法信息，然后在适当的时候复制重放，使系统难辨真伪。例如，C 实体截获了 B 实体发往 A 实体的订单，然后重复地向 A 发送复制的订单，使得 A 的工作出现混乱。

（3）发布或传播不良信息。例如发布垃圾邮件，传播包括色情、暴力、毒品、邪教和赌博等内容的信息。

### 3. 信息系统风险损失

1）信息资源损失

（1）机密性（confidentiality）损失。数据在传输或存储时有被非法截取的可能，就会形成机密性威胁。例如被监听、被分析等。提高信息机密性的方法有数据加密、进行访问控制以及对访问者进行身份验证等，以保证数据不被非授权者知晓。

（2）完整性（integrity）损失。完整性损失是指数据在传输或存储过程中被篡改、被丢失、被破坏的可能。为了保护数据完整性，可以进行数据的完整性校验以及认证等，可以发现数据是否被篡改，进而可以进行数据的恢复。

（3）可用性（availability）损失。可用性损失是指保障合法用户正常使用信息的能力受到的威胁。例如，拒绝访问攻击导致了合法用户正常访问信息资源的能力丧失。

（4）真实性（authenticity）损失。真实性损失主要是指接收方所具有的辨认假冒和抗拒

否认的能力受到威胁。

信息资源损失可以进一步归结为如下 3 类。

（1）信息破坏。非法取得信息的使用权，删除、修改、插入、恶意添加或重发某些数据，以影响正常用户对信息的正常使用。

（2）信息泄密。故意或偶然地非法侦听、截获、分析某些信息系统中的信息，造成系统数据泄密。

（3）假冒或否认。假冒某一可信任方进行通信或者对发送的数据事后予以否认。

2）系统损失

（1）系统被非法访问。

（2）造成通信线路、计算机网络以及主机、光盘、磁盘等系统实体运行不正常或瘫痪，丧失可用性。

**4. 信息系统风险等级**

风险与系统本身的脆弱性（漏洞）的高低和外部威胁的高低有关。也就是说，风险是威胁和漏洞的综合结果。图 1.51 表明了这种关系：风险=威胁+脆弱性，即没有威胁，就不会有风险；同样，没有脆弱性，也不会有风险。

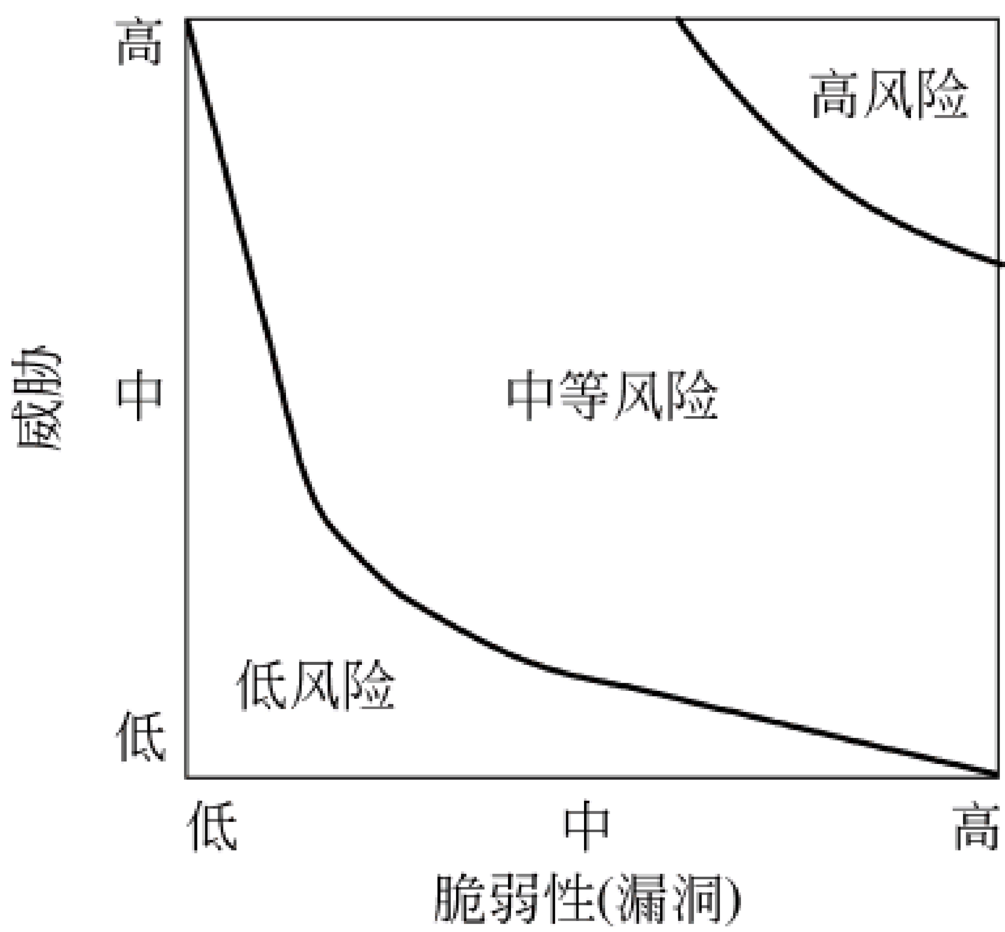


图 1.51　风险与脆弱性-威胁的关系

对威胁和脆弱性进行综合，可以将风险分为低、中、高 3 个级别。

（1）低风险。当系统具有较低的脆弱性或者面临较低的威胁时，其安全将处于低风险级别。

（2）中等风险。当系统具有中等的脆弱性或者面临中等的威胁时，其安全将处于中等风险级别。

（3）高风险。当系统具有较高的脆弱性并且面临较高的威胁时，其安全将处于高风险级别。

## 1.7.2　信息系统安全策略

安全策略是控制和管理主体对客体访问时，为安全目的而制定的一组规则和目标约束，以及为达到安全目的采取的步骤。安全策略可以反映一个组织或一个系统的安全需求。信息安全策略的制定应以信息系统为对象，根据风险分析确立安全方针，并依照这个方针来制定相应的策略。下面是中国信息安全产品评测认证中心提出的系统一般采取的安全策略，供安全管理人员制定系统安全策略时参考。在具体制定系统的安全策略时，可以根据风险分析，从中选择必要的内容，同时根据需求追加一部分内容。

**1. 基于数据资源的安全保护**

1）基于数据资源的安全保护目标

（1）机密性保护，就是保护信息（数据）不被非法泄露或不泄露给那些未授权掌握这

一信息的实体。

（2）完整性保护，就是保护信息（数据）不被未授权的篡改，或被非法篡改后有被恢复的能力。

（3）拒绝否认性（no-repudiation）保护，也称为不可抵赖性或不可否认性保护，即通信双方不能抵赖或否认已经完成的操作或承诺。

2）基于数据资源的安全保护策略

（1）数据隐藏，隐蔽数据的可见性。

（2）数据加密，隐蔽数据的可读性，使攻击者即使获得了数据也难以知道其真实内容。

（3）数字认证。具体手段包括数字签名和数据公证。

（4）数据容错、数据容灾和数据备份。便于数据资源被篡改、被破坏、被丢失后，能及时恢复。

**2. 基于系统的安全保护**

1）基于系统的安全保护目标

（1）可用性（availability）保护，即确保授权用户在需要时可以正确地访问系统中的数据。为此要保证系统能够正常工作，能在确定的条件下、规定的时间内完成规定的功能，实现规定的特性，并使合法用户对于数据资源的使用不会被拒绝。

（2）可控性（controllability）保护，即系统对于信息内容和传输具有控制能力。

（3）有效性（availability）保护，即提供面向用户的服务。简单地说，就是合法者可用，非法者拒绝。

2）基于系统的安全保护

（1）恶意程序的预防、检测和清除。

（2）入侵检测（IDS），当出现不正当访问时，应设置能够将其查出并通知风险管理者的检测功能。

（3）灾害预防与应急处理，具有应对各种灾害的策略和应急处理方案。

（4）防火墙。

（5）授权，包括口令、访问控制、数字证书和身份认证。

（6）日志。

（7）漏洞扫描与渗透测试。

（8）安全审计。

### 1.7.3 信息系统安全防御原则

信息系统的安全是防御式安全。因此在系统的规划、设计、实现、集成、安装和调试等所有过程中，都应同步考虑安全策略和功能具备的程度，坚持预防为主，不要抱有侥幸心理，放掉任何一个已经发现的漏洞和可能的安全威胁。

不同的组织在不同的背景下，基于不同的利益考虑，往往会制定出一些信息系统安全的防御原则。下面介绍目前已经取得共识的信息系统安全防御原则。

**1. 木桶原则**

木桶原则也称为均衡防护原则。它是基于“木桶的容积由其最短的一块木板决定”的原则，考虑到即使绝大部分的环节上防御能力极强，只要有一处很弱，系统总体的防御力也是弱的。因此，要对于最常见的攻击手段采取均衡防护。

**2. 成本效率原则**

任何系统都不是 100%安全的，因为 100%安全要求的成本可能是无限的。因此，成本效率原则要求针对系统的重要性，设定相应的安全需求级别，采取相应的安全措施。

**3. 可扩展性原则**

考虑信息系统的发展、规模的变化以及新的风险的出现，要求安全体系具有可扩展性和延续性。

**4. 分权制衡原则**

要害部位的管理权限不应当交给一个人管理，要将权利分割给几个人，互相制约。

**5. 最小特权原则**

对于某个人只有需要时才可以授予某种特权，但同时要限制其他的系统特权。

**6. 失效保护原则**

系统应当是可控的。一旦系统控制失灵，要有紧急响应预案，阻止风险蔓延和系统恶化。例如，一旦系统发生故障，必须拒绝入侵者的访问；包过滤路由器发生故障，将不允许任何数据包流通；某代理服务器出现故障，就再不提供服务等。

**7. 公开揭露原则**

任何系统遭受某种入侵，都是由于系统存在相应的漏洞。对于发现的漏洞，有人认为应当予以保密，以防更多的入侵。但是，现在普遍的观点是应当公开漏洞，一是可以引起广泛的注意和防护，二是可以发动更多的人或组织提供补救措施，三是可以给入侵者以威慑。

**8. 立足国内原则**

安全技术和设备首先要立足国内，未经批准不得消化、改造或直接应用国外技术和设备。

**9. 可评估原则**

安全系统的规划、设计、实施和运行都要有章可循，同时要考虑用户对于安全的需求和具体环境，使安全系统成为可以论证、可以评估的系统。

# 习　题

## 一、选择题

1. 下列关于计算机病毒的叙述中，________是错误的。

   A. 计算机病毒会造成对计算机文件和数据的破坏

   B. 只要删除感染了病毒的文件就可以彻底消除病毒

   C. 计算机病毒是一段人为制造的小程序

   D. 计算机病毒是可以预防和消除的

2. 计算机病毒是指“能够侵入计算机系统并在计算机系统中，破坏系统正常工作的一种具有繁殖能力的________”。

   A. 一种被破坏了的程序

   B. 具有破坏性，并可以在计算机中潜伏、传播的特殊小程序

   C. 特殊微生物

   D. 源程序

3. 下列关于计算机病毒的说法中，正确的一条是________。

   A. 计算机病毒是一种有损计算机操作人员身体健康的生物病毒

   B. 当U盘或光盘不清洁时，将会传播计算机病毒

   C. 计算机病毒是一种通过自我复制进行传染的，破坏计算机程序和数据的小程序

   D. 计算机病毒是一种有逻辑错误的程序

4. 防止U盘感染病毒的有效方法是________。

   A. 不要把无毒U盘与有毒U盘放在一起

   B. 使U盘写保护

   C. 保持机房清洁

   D. 定期用酒精对U盘进行消毒

5. 计算机宏病毒主要感染________文件。

   A. .EXE　　B. .COM　　C. .TXT　　D. .DOC

6. 计算机病毒最重要的特点是________。

   A. 可执行　　B. 可传染　　C. 可保存　　D. 可打印

7. 为了预防计算机病毒，应采取的正确措施是________。

   A. 每天都对计算机硬盘和软件进行格式化　　B. 不用盗版软件和来历不明的软盘

   C. 不同任何人交流　　D. 不玩任何计算机游戏

8. 下列描述中，不属于引导扇区病毒的是________。

   A. 用自己的代码代替MBR中的代码　　B. 会在操作系统之前加载到内存中

   C. 将自己复制到计算机的每个磁盘　　D. 格式化硬盘

9. 特洛伊木马________。

   A. 表面上看起来无害，但隐藏着罪恶　　B. 不是有意破坏，仅仅制造恶作剧

C. 经常将自己复制，附着在宿主文件中　　D. 传播和运行都不需要客户参与

10. 蠕虫________。

A. 不进行自我复制　　B. 不向其他计算机传播

C. 不需要宿主文件　　D. 不携带有效负载

11. 后门________。

A. 是为计算机系统开启秘密访问入口的程序　　B. 会大量占用计算机资源，造成计算机瘫痪

C. 用于对互联网中的目标主机发起攻击　　D. 用于寻找电子邮件地址，发送垃圾邮件

12. 从安全属性对各种网络攻击进行分类，截获攻击是针对________的攻击。

A. 机密性　　B. 可用性　　C. 完整性　　D. 真实性

13. “会话侦听和劫持技术”是属于________的技术。

A. 密码分析还原　　B. 协议漏洞渗透

C. 应用漏洞分析与渗透　　D. DoS攻击

14. 攻击者用传输数据来冲击网络接口，使服务器过于繁忙以至于不能应答请求的攻击方式是________。

A. 拒绝服务攻击　　B. 地址欺骗攻击

C. 会话劫持　　D. 信号包探测程序攻击

15. 攻击者截获并记录了从A到B的数据，然后从所截获的数据中提取出信息重新发往B，这种攻击称为________。

A. 中间人攻击　　B. 口令猜测器和字典攻击

C. 强力攻击　　D. 回放攻击

16. 拒绝服务攻击的后果是________。

A. 信息不可用　　B. 应用程序不可用

C. 系统死机　　D. 阻止通信

E. 上面几项都是

17. DDoS攻击破坏了信息的________。

A. 可用性　　B. 保密性　　C. 完整性　　D. 真实性

18. 某用户收到一封可疑的电子邮件，要求他提供银行账户及密码。这是一种________攻击手段。

A. 缓存区溢出　　B. 钓鱼　　C. 后门　　D. DDoS

19. 在网络攻击中，攻击者窃取到系统的访问权并盗用资源进行的攻击属于________。

A. 拒绝服务　　B. 侵入攻击　　C. 信息盗窃　　D. 信息篡改

20. 下列关于特征和行为的描述中，不属于DoS的是________。

A. 利用操作系统或应用系统的薄弱环节发起攻击

B. 生成足够多的业务量来拖垮服务器

C. 对计算机系统的资源进行极大的占用

D. 使用电子邮件地址列表来向其他计算机发送自己的副本

## 二、填空题

1. 计算机病毒是一段________程序，它不单独存在，经常是附属在________的起、末端，或磁盘引

导区、分配表等存储器件中。

2. 计算机病毒的5个特征是：主动传染性、破坏性、________、寄生性（隐蔽性）和________。

3. 计算机病毒是________，它能够侵入________，并且能够通过修改其他程序，把自己或者自己的变种复制插入其他程序中；这些程序又可传染别的程序，实现繁殖传播。

4. ________是一组计算机指令或者程序代码，能自我复制，通常嵌入在计算机程序中，能够破坏计算机功能或者毁坏数据，影响计算机的使用。

5. ________是对计算机系统或其他网络设备进行与安全相关的检测，找出安全隐患和可被黑客利用的漏洞。

6. DDoS的攻击形式主要有________和________。

## 三、问答题

1. 举一例说明你所遇到的信息系统安全威胁事件。
2. 什么是0day漏洞？
3. 收集充分的证据，论述病毒程序的特征。
4. 收集资料，解析下列恶意代码的关键技术。

（1）“求职信”病毒。

（2）“主页”病毒。

（3）“欢乐时光”病毒。

（4）“爱虫”病毒。

（5）“美丽杀”病毒。

（6）“万花谷”病毒。

（7）“红色代码”病毒。

5. 在什么情况下，病毒能感染被写保护的文件？
6. 收集资料，解析一种最新病毒的关键技术。
7. 总结现代病毒技术及其发展趋势。
8. 讨论现代病毒检测技术的发展趋势。
9. 讨论现代反病毒技术的发展趋势。
10. 收集资料，讨论针对当前3种流行病毒的查、杀和感染后的恢复方法。
11. 收集资料，讨论针对当前3种流行蠕虫的查、杀和感染后的恢复方法。
12. 分析蠕虫与病毒的区别，收集资料，解析下列蠕虫的关键技术。

（1）蠕虫王。

（2）震荡波。

13. 收集资料，讨论针对当前3种流行木马的防范及清除策略。
14. 什么叫网络木马攻击？
15. 木马有哪些危害？
16. 木马按破坏功能分为哪几种？
17. 典型木马有哪些特性和辅助特点？
18. 木马有哪些植入技术？

19. 木马自动加载方式有哪几种？

20. 木马是如何隐藏自己的？

21. 请举例说明几种常见的木马程序和使用端口号。

22. 木马的远程监控功能有哪些？

23. 如何对木马进行防御？

24. 简述常见的黑客攻击过程。

25. 比较病毒、蠕虫、木马、后门和僵尸。

26. 收集国内外有关病毒和其他恶意程序的网站信息，简要说明各网站的特点。

27. 总结各种电磁波窃听方法的技术要点，提出相应的防范设想。

28. 收集各种手机监听技术要点，提出相应的防范设想。

29. 收集各种网络监听技术要点，提出相应的防范设想。

30. 收集资料，比较下列传输介质上信息被监听的机会和可能。

（1）以太网。

（2）令牌网。

（3）电话网。

（4）有线电视网。

（5）微波和无线电。

31. 请解释以下5种“窃取机密攻击”方式的含义。

（1）网络踩点（Footprinting）。

（2）扫描攻击（Scanning）。

（3）协议栈指纹（Stack Fingerprinting）鉴别（也称操作系统探测）。

（4）信息流嗅探（Sniffering）。

（5）会话劫持（Session Hijacking）。

32. 请解释以下5种“非法访问”攻击方式的含义。

（1）口令破解。

（2）IP欺骗。

（3）DNS欺骗。

（4）重放（Replay）攻击。

（5）特洛伊木马（Trojan Horse）。

33. 分析路由欺骗的原理，并与ARP欺骗和DNS欺骗进行比较。

34. IP欺骗有哪些方法？

35. 请描述局域网间通信时一次完整的ARP欺骗过程。

36. 请描述一次完整的DNS欺骗过程。

37. 简述电子邮件欺骗可能造成的危害有哪些？

38. 试用工具生成一个口令字典。

39. 假定允许使用26个字母和10个数字构造口令，口令长度为6个字符，若采用蛮力攻击，在下列情况下各需要多少时间？

（1）检查一个口令需要0.1s。

（2）检查一个口令需要1μs。

40. 口令破解有哪些方式？口令破解器通常有哪几部分组成？

41. 简述Windows下的口令攻击方法有哪些。

42. 两人试在UNIX系统上进行一次口令攻击对抗。

43. 举例说明黑客是如何进行Web欺骗的？

44. 尽可能多地收集Sniffer产品的数据，进行比较分析，分别指出它们的使用方法和防范措施。

45. 嗅探软件是如何捕捉到数据包并实现数据包过滤功能的？

46. 介绍一种扫描工具的用法，记录扫描结果并对扫描结果进行分析。

47. 请解释全扫描和半扫描的不同。

48. 秘密扫描是如何实现的？

49. 主动协议栈指纹识别OS有哪些方法？

50. 常用的扫描工具有哪些？

51. 如何对扫描进行防御？

52. 什么是缓冲区？什么是缓冲区溢出？

53. 下载一个进行缓冲区溢出攻击的程序，进行分析。

54. 举出几个利用缓冲区溢出进行攻击的病毒名。简述缓冲区溢出攻击的过程。

55. 简述缓冲区溢出攻击的防御方法。

56. 阅读下面的程序，指出其功能。

```
#include <stdio.h>
int main(int argc,char *argv[])
{
   unsigned  char  camary[5];
   unsigned  char  foo[4];
   memset(foo,'\x00',sizeof(foo));
   strcpy(camary,"XXXX");

   fprintf (stderr, "%16u%n%16u%n%32%n%64u%n\n",
      (int *) &foo[0],1,(int*) &foo[1],1 ,(int*) &foo[2],1, (int*) &foo[3]);
   printf ("foo | camary: %02x%02x%02x%02x | %02x%02x%02x%02x\n",
      foo[0],foo[1],foo[2],foo[3],camary[0], camary[1], camary[2], camary[3];
}
```

57. 在实验室中模拟一次SYN Flood攻击的实际过程。

58. 什么是拒绝服务攻击？

59. 简要回答拒绝服务攻击有哪些常用技术？各种技术的特点是什么？

60. 如何防御拒绝服务攻击对系统的危害？

61. 在DDoS攻击中，为什么黑客不直接去控制攻击傀儡机，而要通过控制傀儡机发动进攻呢？

62. 在http://www.fbi.gov/nipc/trinoo.htm上有一个检测和根除trinoo的自动程序。请下载并试用一次。

63. trinoo DDoS有如下一些基本特性，请根据这些特性提出抵御trinoo的策略。

（1）在主控程序与代理程序的所有通信中，trinoo都使用了UDP协议。

（2）Trinoo主控程序的监听端口是27 655，攻击者一般借助Telnet通过TCP连接到主控程序所在计

算机。

（3）所有从主控程序到代理程序的通信都包含字符串"144"，并且被引导到代理的UDP端口27 444。

（4）主控和代理之间的通信受到口令的保护，但是口令不是以加密格式发送的，因此它可以被“嗅探”到并被检测出来。

64. 在网络上下载2～3个DDoS监测软件，安装到自己的计算机上，记录其工作过程。

65. 总结DDoS攻击的防御方法。

66. 浏览3个黑客网站，综述黑客们讨论的热点问题。

67. 信息系统为什么会存在攻击？试从硬件和软件两方面加以回答。

68. 黑客攻击信息系统有哪些方法？

# 第 2 章　数据安全保护

信息系统安全主要是保障信息系统中数据的安全。数据安全防护主要涉及如下内容。

（1）数据的机密性（confidentiality）保护，指保证数据不为非授权者（用户、实体或过程）获取或使用。

（2）数据的完整性（integrity）保护，指保护数据在传输或存储过程中不受到未授权的篡改或破坏。

（3）数据的抗抵赖性（non-repudiation）保护，指在传输数据时必须携带含有自身特质、别人无法复制的信息，防止数据的发送者或接收者事后对自己行为的否认。

本章基于数据加密，讨论数据的机密性、完整性和抗抵赖性保护技术。

## 2.1　数据的机密性保护

数据的机密性保护包括数据的可见性保护和数据的可读性保护。前者可借助于数据隐藏技术，后者可借助于数据加密技术。

### 2.1.1　数据加密基础

数据加密是隐蔽数据的可读性：将可读的数据——明文（plaintext，也称为明码）转换为不可读数据——密文（ciphertext，也称为密码），使非法者不能直接了解数据的内容。加密的逆过程称为解密。

如果用 $P$ 表示明文，用 $C$ 表示密文，则可以将加密写成如下函数形式

$$C = E_{\mathrm{EK}}(P)$$

这里，$E$ 为加密函数，EK 称为加密密钥。

密码系统包括明文空间、密文空间、密钥空间和密码算法 4 个方面。下面通过介绍几种简单的加密方法来介绍加密算法和密钥的概念以及加密算法与密钥之间的关系。

#### 1. 替代密码

替代密码就是将明文中的每个位置的字母都用其他字母代替。比较简单的置换方法是恺撒算法，它将明文中的每个字母都用相隔一定距离的另一个字母代替。例如，将明文 CHINA 中的每个字母都用字母表中后面的距离为 5 的字母代替，就会变成密文 HMNSF。这里“在字母表上移动一个距离”就称为加密算法，距离 5 就称为加密密钥。这种加密的强度非常低，破译者最多只要按字母表试 25 次，就能根据组词规则破译密文。

以法国密码学家 Vigenere 命名的维吉利亚密码就是一种比较复杂的替代密码。设

$$P = \text{data security}，\text{EK=basic}$$

则采用维吉利亚密码的加密过程如下。

（1）制作维吉利亚方阵，如表 2.1 所示。规则是第 *i* 行以 I 打头。

**表 2.1　维吉利亚方阵**

| 明文 | a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| b | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z | A |
| c | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z | A | B |
| ⋮ | | | | | | | | | | | | | | | | | | | | | | | | | | |
| I | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z | A | B | C | D | E | F | G | H |
| ⋮ | | | | | | | | | | | | | | | | | | | | | | | | | | |
| s | S | T | U | V | W | X | Y | Z | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
| t | T | U | V | W | X | Y | Z | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S |
| ⋮ | | | | | | | | | | | | | | | | | | | | | | | | | | |
| z | Z | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y |

（2）按密钥的长度将 *P* 分解为若干节。这里 basic 的长度为 5，故将明文分解为表 2.2 所示的样子。

**表 2.2　按照密钥分解明文**

| 密钥 | b | a | s | i | c |
|---|---|---|---|---|---|
| 明文 | d | a | t | a | s |
| | e | c | u | r | i |
| | t | y | | | |

（3）对每一节明文，利用密钥 basic 进行变换。以明文 d 为例，变化的方法是：由于 d 处于密钥的 b 列，因此在维吉利亚方阵的第 b 行中找到第 d 个字符即是。其他以此类推。于是得到如下密文：

$$C = E_{\mathrm{EK}}(P) = \mathrm{EALIUFCMZKUY}$$

使用 ASCII 码，所发送的位流为

010001010100000010100110001001001010101010100011001000011010011010100101
001001011010101010101011101

这种密码是将明文中的一个字母用一个相应的密文字母替换，称为简单替换密码（simple substitution cipher）或单字母密码（mono alphabetic cipher）。它应用简单，但系统太脆弱，极容易被攻破。于是又设计出多种形式的替换法，例如以下两种替换法。

（1）多名替换密码（homophonic substitution cipher），一个字母可以映射为多个密文字母。例如：

A→5，12，25，56

B→7，17，31，57

⋮

（2）多字母密码（poly alphabetic cipher），字符块被成组加密。例如：

ABA→RTQ

ABB→SLL

⋮

## 2. 换位密码

换位就是将明文中字母的位置重新排列。最简单的换位就是逆序法，即将明文中的字母倒过来输出。例如：

明文：computer system

密文：metsys retupmoc

这种方法太简单，非常容易破解。下面介绍一种稍复杂的换位方法——列换位法。

使用列换位法，首先要将明文排成一个矩阵，然后按列进行输出。为此要解决两个问题。

（1）排成的矩阵的宽度，即矩阵有多少列。

（2）排成矩阵后，各列按什么样的顺序输出。

为此，要引入一个密钥 K，它既可定义矩阵的宽度，又可以定义各列的输出顺序。例如 K=computer，则这个单词的长度（8）就是明文矩阵的宽度，而该密钥中各字母按在字母序中出现的次序，就是输出的列的顺序。表 2.3 为按密钥对明文“WHAT CAN YOU LEARN FROM THIS BOOK”的排列。于是，输出的密文为

WORO NNSX ALMK HUOO TETX YFBX ARIX CAHX

**表 2.3　按密钥排列的明文举例**

| 密钥 | C | O | M | P | U | T | E | R |
|---|---|---|---|---|---|---|---|---|
| 顺序号 | 1 | 4 | 3 | 5 | 8 | 7 | 2 | 6 |
| 明文 | W | H | A | T | C | A | N | Y |
| | O | U | L | E | A | R | N | F |
| | R | O | M | T | H | I | S | B |
| | O | O | K | X | X | X | X | X |

## 3. 简单异或

异或运算具有如下特点：

$$0\oplus 0=0,\ 0\oplus 1=1,\ 1\oplus 0=1,\ 1\oplus 1=0,\ a\oplus a=0,\ a\oplus b\oplus b=a$$

即两个运算数相同，结果为 0；不同，结果为 1。

使用简单异或进行加密，就是将明文与密钥进行异或运算，解密则是对密文用同一密钥进行异或运算，即

$$P\oplus K=C$$

$$C\oplus K=P$$

#### 4. 分组密码

分组密码是一种加密管理方法。它的基本思想是将明文报文编码（例如用 0 和 1 码进行编码），并按照一定的长度（$m$）进行分组，再将各组明文的码分别在密钥的控制下进行加密。例如将明文编码按照 64 位为一组进行分组加密。

采用分组密码的好处是便于标准化，便于在分组（如 x.25 和 IP）网络中被打包传输。其次，由于一个密文组的传输错误不会影响其他密文组，所以容易实现同步。但是，由于相同的密文一定对应相同的明文，所以分组密码不能隐蔽数据模式，同时也不能抵抗组重放、嵌入和删除等攻击。

### 实验 5　加密博弈

#### 1. 实验目的

（1）掌握古典加密程序的设计方法。
（2）掌握进行密码攻击的方法。

#### 2. 实验内容

（1）实验由两（组）人共同完成：一（组）人进行加密程序设计，另一（组）人进行密码攻击程序设计。

（2）程序要求：加密程序由一组程序组成，分别使用置换法、换位法、异或法或它们的组合方法设计。

（3）实验在一些文件上进行。
（4）实验过程中，要记录攻击时间。
（5）一轮实验完成后，加密方与攻击方角色互换，再继续进行另一轮实验。

#### 3. 实验准备

（1）设计实验过程和环境。例如，两方可以通过电子邮件互相传送。
（2）设计加密程序组和攻击程序组。
（3）准备实验用的被加密文件。

#### 4. 推荐的分析讨论内容

分析影响密码加密强度的因素。

### 2.1.2　数据加密体制

#### 1. 对称密码体制和非对称密码体制

如前所述，一个加密过程可以描述为

$$C = E_{\mathrm{EK}}(P)$$

对应地，可以把解密写成

$$P = D_{\mathrm{DK}}(C)$$

其中，$D$ 称为解密函数，DK 为解密密钥。于是，按照 DK 与 EK 的关系，可以分为两种情况。

1）对称密钥体制

若一种加密方法有 DK=EK，则称其为对称密码体制，或称单钥密码，即在这种方法中，加密使用的密钥与解密使用的密钥相同。在对称密码体制中，最为著名的加密算法是 IBM 公司于 1971—1972 年研制成功的数据加密标准（Data Encryption Standard, DES）分组算法，1977 年被定为美国联邦信息标准。

2）非对称密钥体制

若一种加密方法有 DK≠EK，则称其为非对称密码体制，或称双钥密码，即在这种方法中，加密使用的密钥与解密使用的密钥不相同。在非对称密码体制中，最著名的是以 MIT 的 R. Rivest、A. Shamir 和 L. M.Adleman 三位数学家的姓氏首字母命名的 RSA 算法。RSA 加密体制对一对密钥有如下要求。

（1）加密和解密分别用不同的密钥进行，如用加密密钥 EK 对明文 $P$ 加密后，不能再用 EK 对密文进行解密，只能用相应的另一把密钥 DK 进行解密得到明文，即有 $D_{\mathrm{EK}}(E_{\mathrm{EK}}(P))\neq P$ 和 $D_{\mathrm{DK}}(E_{\mathrm{EK}}(P))=P$。

（2）加密密钥和解密密钥可以对调，即 $D_{\mathrm{EK}}(E_{\mathrm{DK}}(P))=P$。

（3）应能在计算机上容易地成对生成，但不能由已知的 DK 导出未知的 EK，也不能由已知的 EK 导出未知的 DK。

### 2. 基于密钥安全的加密

如前所述，密码的安全决定于算法的安全和密钥的安全两个方面。为此在实际中可以采用两种不同的策略：一种称为受限制的算法；另一种称为基于密钥的算法。受限制的算法就是基于算法保密的安全策略。这种策略曾经被使用，但是在现代密码学中已经不再使用。原因如下。

（1）算法是要人掌握的。一旦人员变动，就要更换算法。

（2）算法的开发是非常复杂的。一旦算法泄密，重新开发需要一定的时间。

（3）不便于标准化。由于每个用户单位必须有自己唯一的加密算法，不可能采用统一的硬件和软件产品；否则偷窃者就可以在这些硬件和软件的基础上进行猜测式开发。

（4）不便于质量控制。用户自己开发算法，需要好的密码专家，否则难以保障安全性。

因此，现代密码学认为，所有加密机制的安全性都应当基于密钥的安全性，而不是基于算法实现的安全保密。这就意味着加密算法可以公开，也可以被分析，可以大量生产使用算法的产品，即使攻击者知道了算法也没有关系，只要不知道解密具体使用的密钥，就不能破译密文。因此，保密的关键是保护解密密钥的安全。

**3. 公开密钥体制**

对称密码体制运算效率高，使用方便，加密效率高，是传统企业中最广泛使用的加密技术。但是，由于通信双方使用同样的密钥，因此无论任何一方生成密钥，都要通过一定渠道向对方传送密钥，有可能在传送过程中使密钥泄露，而且通信双方无论任何一方泄密，都会给双方造成损失。由于在非对称密码体制中，加密与解密使用不同的密钥，所以情况大有不同。

1）非对称密钥的生成过程

设通信在 A、B 之间进行，则可以采用下面的方法生成密钥。

（1）A 方产生一对密钥，将其中一个自己保存，另一个传递给 B 方。

（2）B 方也产生一对密钥，将其中一个自己保存，另一个传递给 A 方。

2）非对称密钥体制中的加密方法

在非对称密钥体制中，每端都拥有两个密钥：一个是只有自己知道，其他任何人都不知道的密钥，对于 A 方而言，称为 A 方的私钥，记为 $SK_A$；另一个是对方传来的，自己和对方都知道的密钥，对于 A 方而言，称为 A 方的公钥，记为 $PK_A$。于是非对称密码体制可以提供如图 2.1 所示的方法进行加密。

（1）A 方先用自己的私钥 $SK_A$ 对数据加密，形成密文 $E_{SK_A}(P)$ 再用 B 方的公钥 $PK_B$ 对密文加密，形成双重加密的密文 $E_{PK_B}\left(E_{SK_A}(P)\right)$。

（2）双重加密的密文 $E_{PK_B}\left(E_{SK_A}(P)\right)$ 传送到 B 方后，B 方先用 B 方的私钥 $SK_B$ 进行一次解密，得到 $E_{SK_A}(P)$；再用 A 方的公钥 $PK_A$ 进行二次解密，才能将二重密文最终解密。

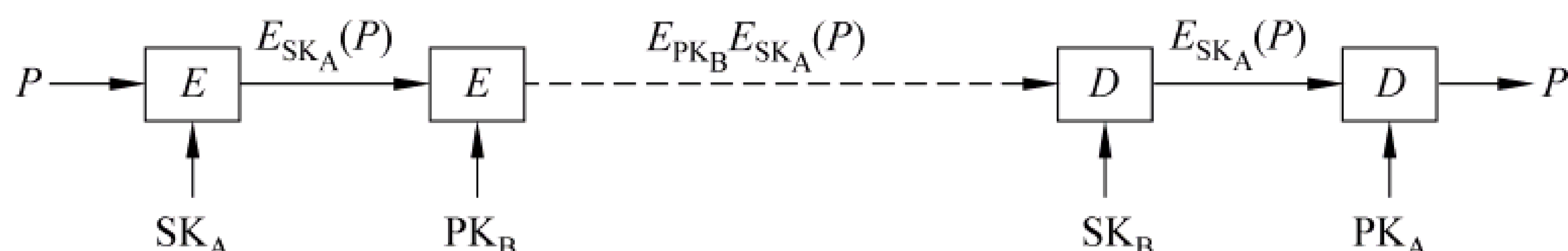



图 2.1　非对称密码体制的加密与解密

在这种情况下，为了保护数据的机密性，只要对每一方的私钥加以保护即可。而对公钥可以不进行保护，甚至可以公开。这样就不存在密钥传输中的失密问题了。因此，通常也将非对称密码体制称为公开密钥体制，因为要向对方传送的一个密钥可以被公开。这也就是通常把公开（非对称）密钥体系中的密钥记为 SK 和 PK，而不再使用记号 EK 和 DK 的原因。

公开密钥体制的问题是算法效率低。因此，一般都是用公开密钥系统传送对称密码体制中的密钥，再用对称密码体制传送密文。

公开密钥体制是斯坦福大学的两位科学家 Diffie 和 Hellman 在 1976 年提出来的。

## 2.1.3　AES 算法

1973 年 5 月 16 日，美国国家标准局（NBS），即现在的美国国家标准与技术研究院

（National Institute of Standards and Technology，NIST），在咨询了 NSA（美国国家安全局）之后，发出通告，公开征求对计算机数据在传输和存储期间进行数据加密的算法。要求如下。

（1）必须提供高度的安全性。

（2）具有相当高的复杂性，使得破译的开销超过获得的利益，但同时又便于理解和掌握。

（3）安全性应当不依赖于算法的保密，加密的安全性仅以加密密钥的保密为基础。

（4）必须适合不同的用户和不同的应用场合。

（5）实现算法的电子器件必须很经济，运行有效。

（6）必须能够有出口。

此后数年内，美国的许多公司、研究机构和大学开发了许多算法。1975 年，IBM 公司提出的算法被采纳，并向全国公布，征求意见。1976 年 11 月，这个算法作为数据加密标准。从此 DES 被广泛应用。但是，由于 DES 的密钥空间较小，只有 56b，所提供的密钥空间只有 $2^{56}$（约为 $7.2\times10^{16}$）。这样的空间不能抵抗穷尽搜索攻击。

为此，1997 年 4 月 15 日，美国 NIST 发起征集新的用于保护敏感的非机密政府信息加密标准——高级加密标准（Advanced Encryption Standard，AES）。1997 年 9 月给出选择 AES 的 3 条评估准则。

（1）安全性。最短密钥长度为 128b，并可抵御各种密钥分析的攻击。

（2）效率。可广泛使用，有很高的计算效率。

（3）灵活性。算法灵活、简洁，能适应各种计算机平台。

1998 年 6 月，NIST 共收到 21 个提交的方案。经过几年的反复论证和评估，最后于 2000 年 10 月 2 日确定选择来自比利时 Katholieke Universiteit Leuven 电子工程系的 Vincent Rijmen 博士和 Proton World International 的 Joan Daemen 博士设计的加密算法 Rijndael（两人的姓氏组合）。在此期间，NIST 宣布 DES 不再作为标准。

Rijndael 算法设计基于非常巧妙的数学原理，经过 AES 标准化后，规定分组大小为 128b，密钥长度可以是 128b、192b 或 256b，分别称为 AES-128、AES-192 和 AES-256。下面介绍 Rijndael 算法。

### 1. 状态矩阵

Rijndael 是分组算法，其运算的基本单位是字节，所给出的分组长度和密钥长度都有 128b（16B）、192b（24B）和 256b（32B）3 个等级。在加密过程中，将每个分组按字节分别组织成如图 2.2 所示的 3 个 4 行字节矩阵。

| | | | |
|---|---|---|---|
| $a_0$ | $a_4$ | $a_8$ | $a_{12}$ |
| $a_1$ | $a_5$ | $a_9$ | $a_{13}$ |
| $a_2$ | $a_6$ | $a_{10}$ | $a_{14}$ |
| $a_3$ | $a_7$ | $a_{11}$ | $a_{15}$ |

(a) 16B矩阵

| | | | | | |
|---|---|---|---|---|---|
| $a_0$ | $a_4$ | $a_8$ | $a_{12}$ | $a_{16}$ | $a_{20}$ |
| $a_1$ | $a_5$ | $a_9$ | $a_{13}$ | $a_{17}$ | $a_{21}$ |
| $a_2$ | $a_6$ | $a_{10}$ | $a_{14}$ | $a_{18}$ | $a_{22}$ |
| $a_3$ | $a_7$ | $a_{11}$ | $a_{15}$ | $a_{19}$ | $a_{23}$ |

(b) 24B矩阵

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $a_0$ | $a_4$ | $a_8$ | $a_{12}$ | $a_{16}$ | $a_{20}$ | $a_{24}$ | $a_{28}$ |
| $a_1$ | $a_5$ | $a_9$ | $a_{13}$ | $a_{17}$ | $a_{21}$ | $a_{25}$ | $a_{29}$ |
| $a_2$ | $a_6$ | $a_{10}$ | $a_{14}$ | $a_{18}$ | $a_{22}$ | $a_{26}$ | $a_{30}$ |
| $a_3$ | $a_7$ | $a_{11}$ | $a_{15}$ | $a_{19}$ | $a_{23}$ | $a_{27}$ | $a_{31}$ |

(c) 32B矩阵

图 2.2　3 个状态矩阵

这些矩阵是 Rijndael 算法进行处理的基础，一个加密过程就是不断对这种矩阵进行迭代

变换的过程，将之称为状态（state）矩阵。状态矩阵的列数记为 $N_b$。同样，也要把 3 个等级的密钥组织成 3 种字节矩阵。

### 2．Rijndael 算法加密的迭代过程

Rijndael 算法是一种迭代算法，其过程如图 2.3 所示。首先将密钥 $K_0$ 和待加密信息按位相与，然后所有要加密的分组都用一个轮函数 $F$ 进行迭代计算。

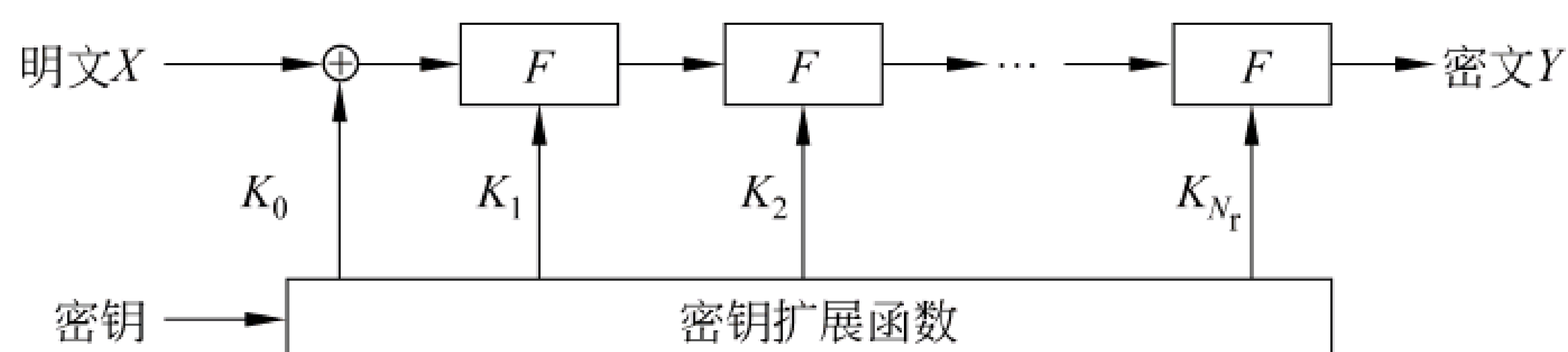



图 2.3　迭代过程

AES 的函数 $F$ 要迭代 $N_r$ 轮，$N_r$ 的值由密钥长度和分组长度决定，具体如表 2.4 所示。

**表 2.4　Rijndael 算法迭代轮数 $N_r$ 的值**

| | 分组长度 128b | 分组长度 192b | 分组长度 256b |
|---|---|---|---|
| 密钥长度 128b（$N_b=4$） | 10 | 12 | 14 |
| 密钥长度 192b（$N_b=6$） | 12 | 12 | 14 |
| 密钥长度 256b（$N_b=8$） | 14 | 14 | 14 |

在前 $N_r-1$ 轮中，$F$ 包含 4 个动作。

（1）字节代换。

（2）行移位。

（3）列混淆。

（4）轮密钥加。

最后一轮（第 $N_r$ 轮）包含 3 个动作。

（1）字节代换。

（2）行移位

（3）轮密钥加。

$K_0, K_1, K_2, \cdots, K_{N_r}$ 为每一轮中使用的子密钥，它们由一个密钥扩展函数所产生。初始密钥 $K_0$ 就是主密钥 $K$。

### 3. 字节代换

在 Rijndael 算法中采用了替代技术，进行字节代换。字节代换是非线性的，它独立地将状态中的每个字节用代换表——S 盒中的值进行代换。如图 2.4 所示，S 盒用每个字节的高 4 位作为行值，低 4 位作为列值，按此找出对应的表值。表中的数值都是十六进制的。

### 4. 行移位

在 Rijndael 算法中还采用了换位技术——行移位。移动方法是对状态阵列中的第 2、3、

| 行＼列 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 63 | 7C | 77 | 7B | F2 | 6B | 6F | C5 | 30 | 01 | 67 | 2B | FE | D7 | AB | 76 |
| 1 | CA | 82 | C9 | 7D | FA | 59 | 47 | F0 | AD | D4 | A2 | AF | 9C | A4 | 72 | C0 |
| 2 | B7 | FD | 93 | 26 | 36 | 3F | F7 | CC | 34 | A5 | E5 | F1 | 71 | D8 | 31 | 15 |
| 3 | 04 | C7 | 23 | C3 | 18 | 96 | 05 | 9A | 07 | 12 | 80 | E2 | EB | 27 | B2 | 75 |
| 4 | 09 | 83 | 2C | 1A | 1B | 6E | 5A | A0 | 52 | 3B | D6 | B3 | 29 | E3 | 2F | 84 |
| 5 | 53 | D1 | 00 | ED | 20 | FC | B1 | 5B | 6A | CB | BE | 39 | 4A | 4C | 58 | CF |
| 6 | D0 | EF | AA | FB | 43 | 4D | 33 | 85 | 45 | F9 | 02 | 7F | 50 | 3C | 9F | A8 |
| 7 | 51 | A3 | 40 | 8F | 92 | 9D | 38 | F5 | BC | B6 | DA | 21 | 10 | FF | F3 | D2 |
| 8 | CD | 0C | 13 | EC | 5F | 97 | 44 | 17 | C4 | A7 | 7E | 3D | 64 | 5D | 19 | 73 |
| 9 | 60 | 81 | 4F | DC | 22 | 2A | 90 | 88 | 46 | EE | B8 | 14 | DE | 5E | 0B | DB |
| A | E0 | 32 | 3A | 0A | 49 | 06 | 24 | 5C | C2 | D3 | AC | 62 | 91 | 95 | E4 | 79 |
| B | E7 | C8 | 37 | 6D | 8D | D5 | 4E | A9 | 6C | 56 | F4 | EA | 65 | 7A | AE | 08 |
| C | BA | 78 | 25 | 2E | 1C | A6 | B4 | C6 | E8 | DD | 74 | 1F | 4B | BD | 8D | 8A |
| D | 70 | 3E | B5 | 66 | 48 | 03 | F6 | 0E | 61 | 35 | 57 | B9 | 86 | C1 | 1D | 9E |
| E | E1 | F8 | 98 | 11 | 69 | D9 | 8E | 94 | 9B | 1E | 87 | E9 | CE | 55 | 28 | DF |
| F | 8C | A1 | 89 | 0D | BF | E6 | 42 | 68 | 41 | 99 | 2D | 0F | B0 | 54 | BB | 16 |

图 2.4　Rijndael 算法中使用的 S 盒

4 行分别循环左移 $C_1$、$C_2$、$C_3$ 列。$C_1$、$C_2$、$C_3$ 的值与分组大小有关，具体值见表 2.5。

**表 2.5　状态矩阵中第 2、3、4 行的移位值**

| $N_b$ | $C_1$ | $C_2$ | $C_3$ |
|---|---|---|---|
| 4 | 1 | 2 | 3 |
| 6 | 1 | 2 | 3 |
| 8 | 1 | 3 | 4 |

图 2.5 为对分组长度为 128b 的状态矩阵列行移位的情形。

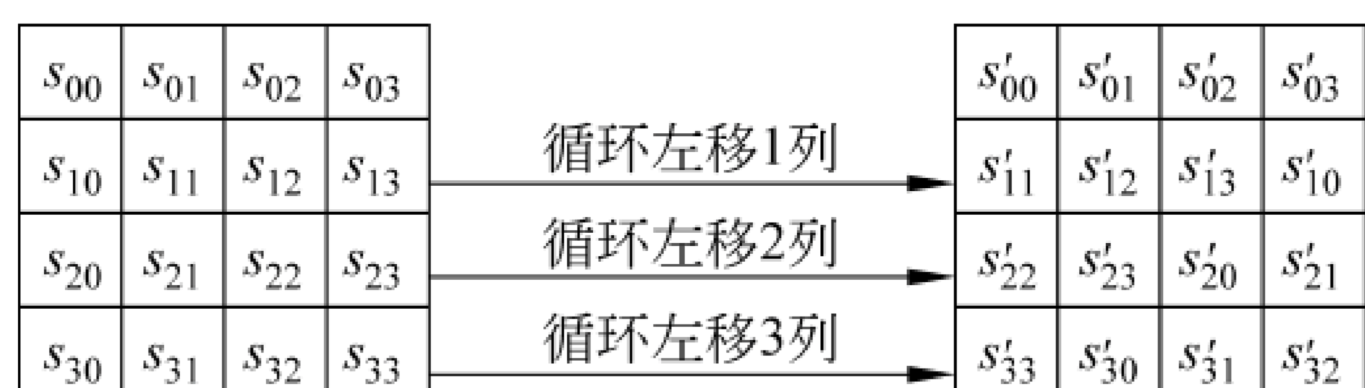



图 2.5　分组长度为 128b 的状态矩阵列行移位

**5. 列混淆**

在行位移的基础上，Rijndael 算法还进行了列混淆，即对状态矩阵中的每一列（在 Rijndael 算法中称为一个字）进行一次如下的矩阵运算。其中 $c$ 为列号（$0 \leqslant c < N_b$），如 $2c$ 为第 2 行第 $c$ 列。

$$\begin{bmatrix} s'_{0c} \\ s'_{1c} \\ s'_{2c} \\ s'_{3c} \end{bmatrix} = \begin{bmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{bmatrix} \begin{bmatrix} s_{0c} \\ s_{1c} \\ s_{2c} \\ s_{3c} \end{bmatrix}$$

即有如下结果：

$$s'_{0c} = (\{02\} \bullet s_{0c}) \oplus (\{03\} \bullet s_{1c}) \oplus s_{2c} \oplus s_{3c}$$
$$s'_{1c} = s_{0c} \oplus (\{02\} \bullet s_{1c}) \oplus (\{03\} \bullet s_{2c}) \oplus s_{3c}$$
$$s'_{2c} = s_{0c} \oplus (s_{1c} \oplus (\{02\} \bullet s_{2c}) \oplus (\{03\} \bullet s_{3c})$$
$$s'_{3c} = (\{03\} \bullet s_{0c}) \oplus s_{1c} (\{03\} \bullet s_{1c}) \oplus s_{2c} \oplus (\{02\} \bullet s_{3c})$$

这里，符号⊕表示二进制异或。

**例 2.1** 若第 1 列分别为{87}、{6E}、{46}、{A6}，请计算 $s'_{01}$。

由上述混淆算法可以得到

$$s'_{01} = (\{02\} \bullet \{87\}) \oplus (\{03\} \bullet \{6E\}) \oplus \{46\} \oplus \{47\}$$

用多项式表示为

$$\{02\} = x$$
$$\{87\} = x^7 + x^2 + x + 1$$

则

$$\{02\} \bullet \{87\} = x^8 + x^3 + x^2 + x$$

再对一个 8 次的不可约多项式求模，得

$$(x^8 + x^3 + x^2 + x) \bmod (x^8 + x^4 + x^3 + x^2 + x + 1) = x^4 + x^2 + 1$$

写成二进制形式为 00010101。

同理得到{03} • {6E}= 10110010，{46}=01000110，{47}=10100110。

故表达式（{02} • {87}）⊕（{03} • {6E}）⊕ {46} ⊕ {47}可以用下面的方法计算得到。

```
    0001 0101
    1011 0010
    0100 0110
⊕   1010 0110
-------------
    0100 0111={47}
```

**6. 轮密钥加**

轮密钥加变换可以看成状态矩阵中的一个字与轮密钥的一个字进行异或运算。

**7. 密钥扩展**

下面以分组长度和密钥长度都为 128b 的 AES 加密算法为例，说明密钥扩展的算法。由前面的讨论可知，这时要进行 10 轮迭代，总共需要 $K_0, K_1, K_2, \cdots, K_{N_r}$ 共 11 个密钥。每个密钥矩阵有 4 个字，共需要 44 个密钥字；而现在只有 4 个字的原始密钥。Rijndael 算法中的密钥扩展就是要从这 4 个原始的密钥字再扩展出来 40 密钥字。密钥扩展算法如下。

首先建立一个大小为 44 的密钥数组 $w$。其每个元素为 4 个字节。接着将原始密钥复制到 $w$ 的前 4 个字中，得到 $w[0]$、$w[1]$、$w[2]$、$w[3]$。然后用这 4 个字扩展 $w$ 中余下的部分，使 $w[i]$的值依赖于 $w[i-1]$和 $w[i-4]$。其中 $i \geqslant 4$。然后按照下列方法进行扩展。

（1）当 $i$ 为非 4 的整数倍时，$w[i]$的值为 $w[i-1]$与 $w[i-4]$的异或。

（2）当 $i$ 为 4 的整数倍时，按下面的方法进行。

① 将 $w[i-1]$的 4 个字节$[b_0, b_1, b_2, b_3]$循环左移 1 个字节，即变为$[b_1, b_2, b_3, b_0]$。

② 用 S 盒对输入字的每个字节进行字节代换。

③ 将 $w[i-4]$与①的结果异或，与②的结果异或，再与轮常数 Rcon[$i$]异或。轮常数在每一轮中为一个常数，具体见表 2.6 所示。

**表 2.6　轮常数的值**

| 轮数 $i$ | Rcon[$i$] | 轮数 $i$ | Rcon[$i$] |
|---|---|---|---|
| 1 | 010000000 | 6 | 200000000 |
| 2 | 020000000 | 7 | 400000000 |
| 3 | 040000000 | 8 | 800000000 |
| 4 | 080000000 | 9 | 1B0000000 |
| 5 | 100000000 | 10 | 360000000 |

192b 和 256b 中的密钥扩展算法如上类似，在此不再赘述。

#### 8．Rijndael 解密算法

Rijndael 解密算法是 Rijndael 加密算法的逆变换，算法类似，只是顺序不同，这里不再介绍。

### 2.1.4　公开密钥算法 RSA

#### 1．RSA 的数学基础

1）费马（Fermat）定理

描述 1：若 $p$ 是素数，$a$ 是正整数且不能被 $p$ 整除，则 $a^{p-1} \equiv 1 \bmod p$。

描述 2：对于素数 $p$，若 $a$ 是任一正整数，则 $a^p \equiv a \pmod p$。

**例 2.2**　设 $p=3$，$a=2$，则 $2^{3-1}=4 \equiv 1 \pmod 3$ 或 $2^3=8 \equiv 2 \pmod 3$。

**例 2.3**　设 $p=5$，$a=3$，则 $3^{5-1}=81 \equiv 1 \pmod 5$ 或 $3^5=243 \equiv 3 \pmod 5$。

2）欧拉（Euler）函数

欧拉函数 $\varphi(n)$表示小于 $n$ 并与 $n$ 互素的正整数的个数。

**例 2.4**　$\varphi(6)=2, \{1,5\}$；$\varphi(7)=6, \{1, 2, 3, 4, 5, 6\}$；$\varphi(9)=6$，$\{1, 2, 4, 5, 7, 8\}$。

3）欧拉定理

若整数 $a$ 和 $m$ 互素，则

$$a^{\varphi(m)} \equiv 1 \pmod m$$

**例 2.5** 设 $a$=3，$m$=7，则有 $\varphi(7)=6$，$3^6=729$，$729\equiv 1$（mod 7）。

**例 2.6** 设 $a$=4，$m$=5，则有 $\varphi(5)=4$，$4^4=256$，$256\equiv 1$（mod 5）。

### 2．RSA 加密密钥的产生

RSA 依赖于一个基本假设：分解因子问题是计算上的困难问题，即很容易将两个素数乘起来，但分解该乘积是困难的。

1）基本过程

（1）选取两个保密的大素数 $p$ 和 $q$（保密）。

（2）计算 $n=pq$（公开），$\varphi(n)=(p-1)(q-1)$（保密）。

（3）随机选取一个整数 $e$，满足 $1<e<\varphi(n)$且 $\gcd(\varphi(n),e)=1$（公开）。

（4）计算 $d$，满足 $de\equiv 1 \bmod \varphi(n)$（保密）。

说明：$d$ 是 $e$ 在模 $\varphi(n)$下的乘法逆元。因为 $e$ 与 $\varphi(n)$互素，所以其乘法逆元一定存在。

（5）得到一对密钥：公开密钥$\{e, n\}$，秘密密钥$\{d, n\}$。

2）应用举例

（1）选择两个素数 $p=7$，$q=17$。

（2）计算 $n=pq=7\times17=119$。

计算 $n$ 的欧拉函数 $\varphi(n)=(p-1)(q-1)=6\times16=96$。

（3）从[0，95]间选一个与 96 互质的数 $e=5$。

（4）根据式

$$5d=1 \bmod 96$$

解出 $d=77$，因为 $ed=5\times77=385=4\times96+1\equiv 1 \bmod 96$。

（5）得到公钥 PK=$(e, n)=\{5,119\}$，密钥 SK=$\{77,119\}$。

### 3．RSA 加密/解密过程

1）基本过程

（1）明文数字化，即将明文转换成数字串。

（2）分组。将二进制的明文串分成长度小于 $\log_2 n$ 的数字分组。如果 $p$ 和 $q$ 都为 100 位素数，则 $n$ 将有 200 位，所以每个明文分组应小于 200 位。

（3）加密算法

$$C_i=M_i^e \bmod n$$

最后得到的密文 $C$ 由长度相同的分组 $C_i$ 组成。

（4）解密算法

$$D(C_i)\equiv C_i^d \bmod n$$

2）综合应用举例

（1）产生密钥。

设 $p=43, q=59, n=43\times59=2537$，$\varphi(n)=42\times58=2436$。

取 $e$=13（与 $\varphi(n)$没有公因子）。

解方程 $de\equiv1$（mod 2436），计算过程如下：

$$
\begin{aligned}
2436 &= 13\times187+5，\quad 5=2436-13\times187\\
13 &= 2\times5+3，\quad 3=13-2\times5\\
1 &= 3-2=3-(5-3)=2\times3-5=2\times(13-2\times5)-5\\
&= 2\times13-5\times5\\
&= 2\times13-5\times(2436-13\times187)\\
&= (187\times5+2)\times13-5\times2436\\
&= 937\times13-5\times2436
\end{aligned}
$$

即

$$937\times13\equiv1\ (\text{mod } 2436)$$

故 $e$=13，$d$=937。

（2）加密。

明文：public key encryptions。

明文分组：pu bl ic ke ye nc ry pt io ns。

明文数字化（按字母序，令 $a=00, b=01, c=02,\cdots,\ y=24, z=25$）：

1520 0111 0802 1004 2404 1302 1724 1519 0814 1418

加密：按照算法 $M_i^e\ (\text{mod } n)=C_i$，如 $1520^{13}\ (\text{mod } 2537)=0095$，得到密文：

0095 1648 1410 1299 1365 1379 2333 2132 1751 1289

解密：按照算法 $C_i^d\ (\text{mod } n)=M_i$，如 $0095^{937}\ (\text{mod } 2537)=1520$。

#### 4．RSA 安全性分析

RSA 体制的加密强度依赖于大数分解的困难程度。采用穷举法，对于两个 100 位的十进制大素数，破译它大约需要 $10^{23}$ 步，若使用 100 万步/秒的计算机资源对其进行破密，约需要 1000 年。

但是，人类的计算能力也在不断提高，原来一些被认为不可能分解的大数，现在已经被成功分解。例如，RSA-129（即 $n$ 为 129 位的十进制数，约 428b），历时 8 个月，已经于 1994 年 4 月被成功分解。而且有报道，国外科学家正在用量子方法对大数分解发起冲击。

不过，在目前的情况下，密钥长度在 1024～2048b 的 RSA 还是相对安全的。

为了保证 RSA 安全性，对 $p$ 和 $q$ 还有如下要求。

（1）$p$ 和 $q$ 的长度相差不要太大。

（2）$p-1$ 和 $q-1$ 都应当有大数因子。

（3）gcd($p-1$，$q-1$)应小。

### 2.1.5　密钥管理

现代密码体制有点像门锁，房间的安全主要依赖于钥匙。也更像现代的电子锁，锁门的操作（相当于加密算法）非常简单，房间的安全关键在于卡片式钥匙的保管，而一旦卡片丢失或房间换人，只要重新对卡片进行设置即可。按照“一切秘密寓于密钥之中”的现

代密码学基本原则，密钥的安全保护成为系统安全的一个重要方面，密钥管理是系统安全中的一件至关重要的工作。

### 1. 密钥管理的一般过程

一般说来，密钥管理主要包括密钥的生成、分配、使用等一系列过程。下面简要介绍这些过程。

1）密钥的生成

密钥生成的目的是生成好的密钥。对于对称加密来说，密钥的长度越大，对应的密钥空间就越大，密钥的强度就大。此外，由自动密钥设备生成的随机比特串要比按照某种规则生成的密钥好。但是，在选择随机生成的密钥时，要避免选择弱密钥。对于公钥密码体制来说，密钥还必须满足特定的数学特征。

2）密钥的分配

在密钥管理中，最核心、最关键的问题是密钥分配——主要涉及密钥的发送和验证。前者要求通过非常安全的通路进行传送，后者要求有一套机制用于检验分发和传送的正确性。

密钥的分发方法可以分为两种：网外分发和网内分发。网外分发即人工分发：派非常可靠的信使（邮寄、信鸽等）携带密钥分配给各用户。但是，随着用户的增加、通信量的增大以及黑客技术的发展，密钥的使用量增大，且要求频繁更换，信使分配就不再适用，而多采用网内密钥分配，即自动密钥分配。

网内密钥分配的方式有两种：用户之间直接分配和通过设立一个密钥分配中心（Key Distribution Center，KDC）分配。具体过程由密钥分配协议决定。目前国际有关标准化机构都在着手制定关于密钥管理技术的规范。

3）密钥的控制使用

控制密钥使用，是为了保证按照预定的方式使用。控制密钥使用的信息有以下几项。

（1）密钥主权人。

（2）密钥合法使用期限。

（3）密钥标识符。

（4）密钥预定用途。

（5）密钥预定算法。

（6）密钥预定使用系统。

（7）密钥授权用户。

（8）在密钥生成、注册、证书等有关实体中的名字等。

4）密钥的保护与存储

密钥从产生到终结，在整个的生存期中都需要保护。一些基本的措施如下。

（1）密钥绝不能以明文形式存放。

（2）密钥首先选择物理上最安全的地方存放。

（3）在有些系统中可以使用密钥碾碎技术由一个短语生成单钥密钥。

（4）可以将密钥分开存放。例如，将密钥平分成两段，一段存入终端，一段存入 ROM；或者将密钥分成若干片，分发给不同的可信者保管。

5）密钥的停用和更新

任何密钥都不可能无限期地使用。有许多因素使得密钥不能使用太长的时间，密钥使用得越久，攻击者对它的攻击方法越多，攻击的机会越多。密钥一旦泄露，若不立即废除，时间越长，损失越大。因此，不同的密钥应当有不同的有效期，同时必须制定一个检测密钥有限期的策略。密钥的有限期依据数据的价值和给定时间里加密数据的数量确定。

当发生下列情况时，应当停止密钥的使用，更新密钥。

（1）密钥的使用期到，应该更新密钥。

（2）确信或怀疑密钥被泄露，密钥及其所有变形都要替换。

（3）怀疑密钥是由一个密钥加密密钥或其他密钥推导出来时，各层与之相关的密钥都应更换。

（4）通过对加密数据的攻击可以确定密钥时，在这段时间内必须更换密钥。

（5）确信或怀疑密钥被非法替换时，该密钥和相关密钥都要被更换。

6）密钥的销毁

密钥被替换后，旧密钥必须销毁。旧密钥虽然不再使用，却可以给攻击者提供许多有重大参考价值的信息，为攻击者推测新的密钥提供许多有价值的信息。为此，必须保证被销毁的密钥不能给任何人提供丝毫有价值的信息。下面是在销毁密钥时使用的一些方法。

（1）密钥写在纸上时，要把纸张切碎或烧毁。

（2）密钥存在 $E^2$PROM 中时，要对 $E^2$PROM 进行多次重写。

（3）密钥存在 EPROM 或 PROM 中时，应将 EPROM 或 PROM 打碎成小片。

（4）密钥存在磁盘中时，应当多次重写覆盖密钥的存储位置，或将磁盘切碎。

（5）要特别注意对存放在多个地方的密钥的同时销毁。

**2. 密钥分配方法举例**

密钥分配是密钥管理的核心。下面介绍几种密钥分配方法。

1）密钥分配中心分发单密钥

若 A 和 B 有一个共同的可信任的第三方——密钥分配中心（KDC）。KDC 可以通过加密连接将密钥安全地传送给 A 和 B。这种方法多用于单钥密钥的分配。图 2.6 为一个采用这

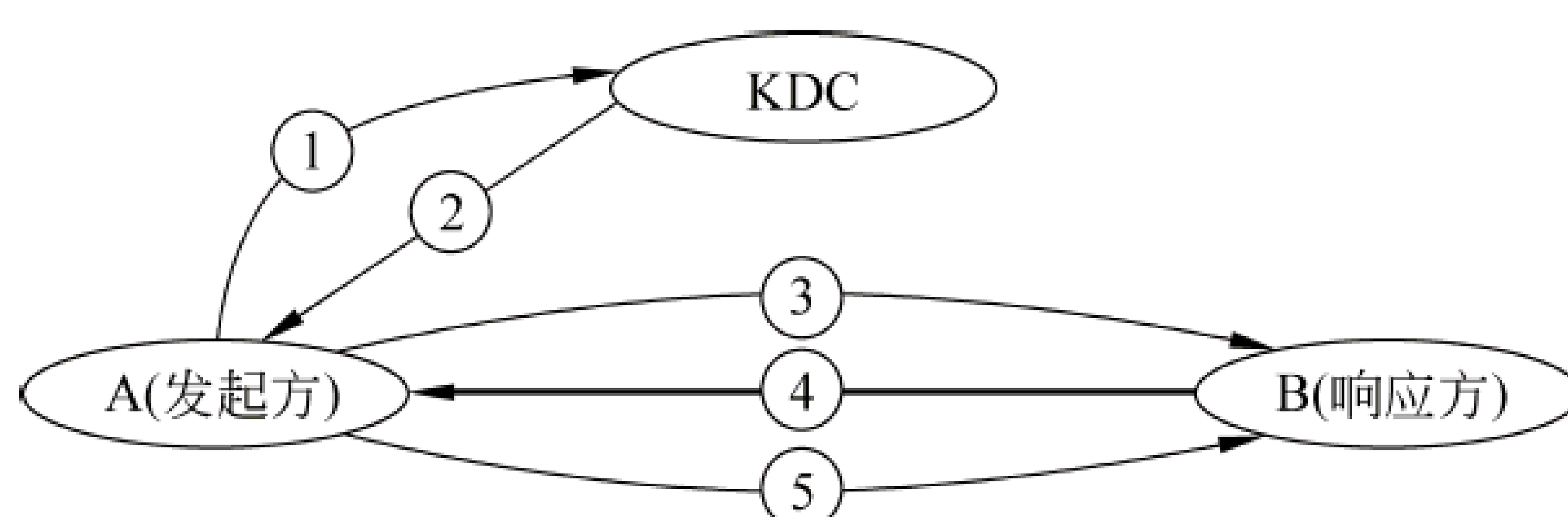



图 2.6　一个依靠 KDC 进行密钥分配的例子

种方法的密钥分配例子。

这个例子的前提是 A 和 B 有一个共同的可信任的第三方——KDC，即 KDC 分别与 A 和 B 有一个保密的信道，即 KDC 与 A 和 B 分别已经有一个通信密钥 $K_A$ 和 $K_B$。假定 A 与 B 的通信是 A 主动，目的是通过 KDC 分配的密钥与 B 建立一个秘密通信通道。过程如下。

（1）A 向 KDC 发出会话密钥请求：$ID_A \| ID_B \| N_A$。

① $ID_A$ 和 $ID_B$ 标识会话双方 A 和 B。

② $N_A$ 标识本次会话（可能是时间戳或随机数等一个他人难于猜测的现时值）。

（2）KDC 对 A 的请求应答：$E_{K_A}[K_s \| ID_B \| N_A \| E_{K_B}[\{K_s \| ID_A\}]]$。

全部报文用 A 已经掌握的密钥 $K_A$ 加密，内容包括 3 部分。

① 一次性会话密钥 $Ks$。

② A 的请求报文（供 A 检验）。

③ 要求 A 中转，但 A 不能知道内容的、用 $K_B$ 加密的一段报文：$K_s \| ID_A$。

（3）A 存储 $K_s$，并向 B 转发 $E_{K_B}[K_s \| ID_A]$。B 得到：

① $K_s$，还知道 $K_s$ 来自 KDC（因为用 $K_B$ 可解密，而 A 不知道 $K_B$，只有 KDC 知道 $K_B$）。

② 由 $ID_A$ 知道会话方是 A。

（4）B 向 A 回送报文：$E_{K_s}[N_B]$。

① 用 $K_s$ 表明自己的身份是 B（因为 $K_s$ 要用 $K_B$ 解密）。

② 用 $N_B$ 防止回放攻击。

（5）A 向 B 回送报文：$E_{K_s}[f(N_B)]$。确认 B 前次收到的报文不是回放。

这样，A 与 B 就有了自己的秘密通道了。

2）无中心的单钥分配

图 2.7 是在无 KDC 或不依靠 KDC 时 A、B 两方建立会话密钥的过程。

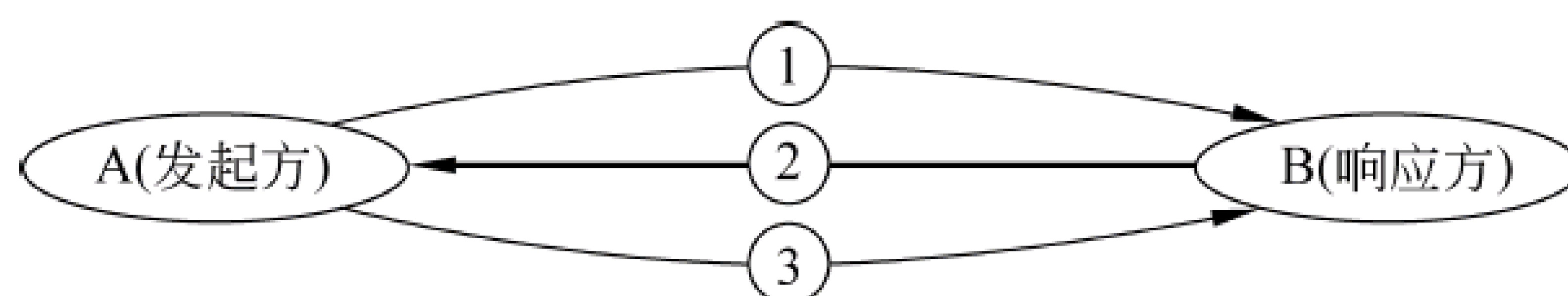



图 2.7　一个无中心的单密钥分配例子

（1）A 向 B 发出会话请求：$N_A$。

$N_A$ 标识本次会话（可能是时间戳或随机数等一个他人难于猜测的现时值）。

（2）B 对 A 的请求应答：$E_{K_M}[K_s \| ID_B \| f(N_A) \| N_B]$。

全部报文用 A、B 共享的主密钥 $K_M$ 加密，内容包括 4 部分。

① B 选取的会话密钥 $K_s$。

② A 的请求报文（包括 $f(N_A)$，供 A 检验）。

③ B 的身份 $ID_B$。

④ 标识本次会话的 $N_B$。

（3）A 存储 $K_s$，并向 B 返回用 $K_s$ 加密的 $f(N_2)$，供 B 检验。

采用这种密钥分配方法，在每一对通信主体之间都需要一个共享主密钥。对于一个有 $n$

个通信主体的网络，主密钥的数量达到 $n(n-1)/2$ 个。当网络较大时，这种方法没有什么实用价值。而依靠 KDC 进行密钥分发仅需要 $n$ 个（KDC 与每个通信实体之间共享的）主密钥。

3）由公钥管理机构分发公钥

若存在一个公钥管理机构，并且所有用户都知道该公钥管理机构的公钥，而只有该公钥管理机构知道自己的私钥，则可以采用图 2.8 所示的方法进行 A、B 公开密钥体制的密钥分配。

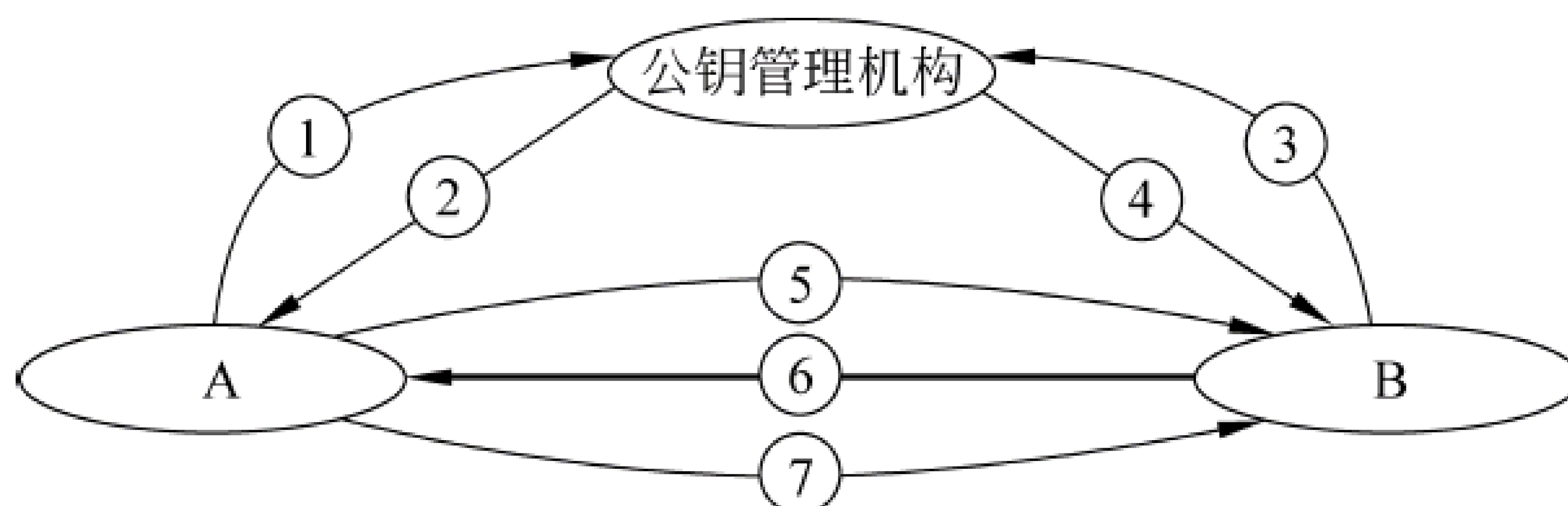



图 2.8　一个依靠公钥管理机构进行密钥分配的例子

这个过程分为 7 步。

① 用户 A 向公钥管理机构发出请求报文，内容如下。

- 一个带时间戳的报文。
- 请求获取 B 的公钥的请求。

② 公钥管理机构对 A 应答（用 A 的公钥加密，A 用自己的私钥解密）。内容有：

- B 的公钥 $PK_B$（供 A 向 B 发加密报文）。
- A 的请求（供 A 验证本报文是对自己请求的应答）。
- 最初的时间戳（供 A 确认不是公钥管理机构发来的旧报文，以确定 $PK_B$ 是 B 的）。

③、④ B 用与①、②相同的方法，从公钥管理机构得到 A 的公钥 $PK_A$。

⑤ A 用 $PK_B$ 向 B 发送一个报文，内容如下。

- A 的身份 $ID_A$。
- 一次性随机数 $N_A$。

⑥ B 用 $PK_A$ 向 A 发送一个报文，内容如下。

- $N_A$（由于只有 B 才能解密用 $PK_B$ 加密的报文，将 $N_A$ 返回 A，让 A 确认是 B）。
- 一次性随机数 $N_B$。

⑦ A 用 $PK_B$ 将 $N_B$ 加密，返回 B，供 B 确认。

这种方法是基于公钥目录表的。公钥目录表是由某个可信的公钥管理机构管理并定期更新、定期公布的用户公钥目录表。目录表中的每个目录项由两个数据组成：用户名和该用户的公钥。

用户可以在公钥目录表的管理机构注册自己的公钥，也可以随时更换现有的公钥，还可以通过电子方式在有安全认证的情况下访问公钥目录表。

应当注意的是，公钥目录表可能会被攻击者伪造、监听、攻击。

4）公钥证书

公钥证书是由证书授权中心或认证中心（Certificate Authority，CA）为用户发布的一种电子证书。例如，用户 A 的证书内容形式为

$$C_A = E_{SK_{CA}}[T, ID_A, PK_A]$$

其中：

① $ID_A$ 是用户 A 的标识。

② $PK_A$ 是 A 的公钥。

③ $T$ 是当前时间戳，用于表明证书的新鲜性，防止发送方或攻击者重发一个旧证书。

④ $SK_{CA}$ 是 CA 的私钥。证书是用 CA 的私钥加密的，以便让任何用户都可以解密，并确认证书的颁发者。

当一方要与另一方建立保密信道时，就要把自己的证书发给对方。接收方用 CA 的公钥对证书进行查验，可以获得发送方的公钥。接收方同意进行保密通信，也要将自己的证书发送给对方。这样，就不依赖 CA 而直接交换了公钥。

## 2.1.6 流密码

### 1. 流密码概述

如前所述，一个加密体系的安全性主要系于密钥上。为了提高安全性，人们曾经致力于选取尽可能长的密钥。但是，长密钥的存储和分配都很困难。于是转而提倡不断改变密钥。早在 1949 年，Shannon 就已经证明：“一次一密”的密码体制是绝对安全的。

流密码（stream cipher）也称为序列密码，就是在“一次一密”的追求中发展起来的一种密码技术。那么，如何实现“一次一密”呢？人们可以从随机数序列的产生中得到启发。人们知道，对于一个随机数发生器，当对其输入不同的随机数种子时，就会生成不同的随机数序列。按照这一原理，对一个密钥流发生器输入不同的种子密钥，也就会生成不同的密钥序列。

流密码以一位或一个字节为单位，使用密钥流中的随机密钥对明文进行加密。例如，密钥流中的一个字节 01001010 对明文流中的一个字节 10010011 进行异或，就可得到密文流中的一个字节 11011001。同样，将密文流与密钥流异或，就可以得到明文流。显然这是一种对称加密技术。

### 2. 同步流密码与自同步流密码

在流密码技术中，如果密钥流完全独立于明文流或密文流，则称这种流密码为同步流密码（synchronous stream cipher）；如果密钥流的产生与明文流或密文流有关，则称这种流密码为自同步流密码（self-synchronous stream cipher）。图 2.9 为同步流密码与自同步流密码示意图。

所谓“同步”，是指在保密通信过程中，通信的双方必须保持精确的同步，接收方才能正确解密。如果通信过程中丢失或增添了一个密文字符，接收方就不能正确解密，直到同步恢复。这种现象称为同步密钥流对于失步的敏感性。这种敏感性是一个缺点，但也带来

一个优点：可以容易地检测到插入、删除、重播等攻击。另一方面，同步流密码的各加密单元（位或字节）之间相对独立，互不相关，不会形成错误传播。

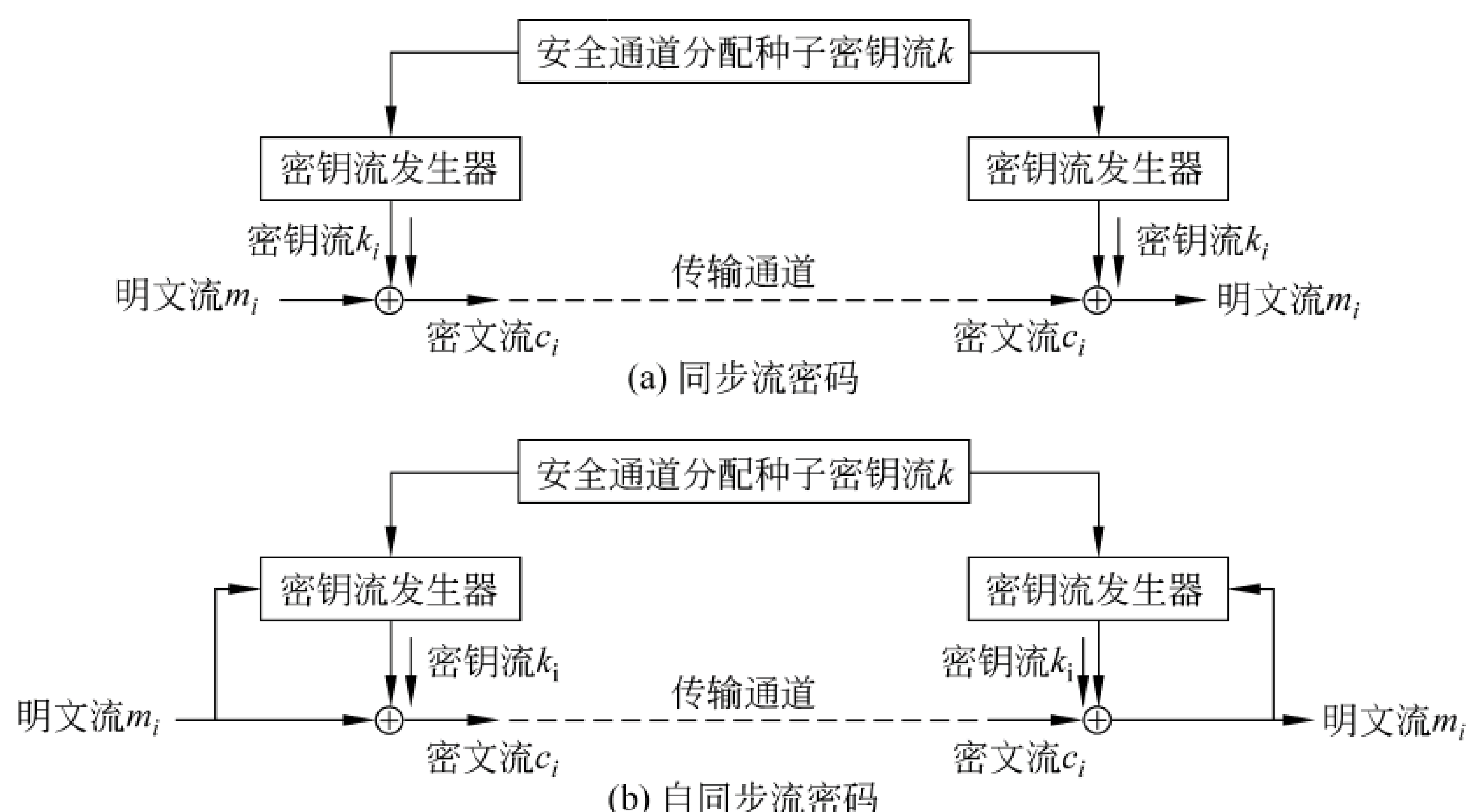



图 2.9　同步流密码与自同步流密码的比较

自同步流密码技术是让每一个加密单元（位或字节）的密钥除了依赖于种子密钥，还依赖于前面 $n$ 个明文（或密文）加密单元。采用这种密码技术，如果在传输中丢失或更改了一个加密单元，将会造成这一错误向后传播 $n$ 个加密单元。不过，在收到 $n$ 个正确的加密单元后，密钥流又会自动同步。

### 3. 密钥流发生器

流密码的安全性完全取决于它的密钥流发生器所产生的密钥流的特性。可以想象，如果一个密钥流是无限长并且是无周期性的真随机序列，则才是真正的“一次一密”的密码体制。但遗憾的是，任何人为地产生的随机序列都不可能达到这样的条件要求。人工只能产生伪随机序列：长度有限。但是人工产生的伪随机序列接近真正随机序列的程度有所不同。这种不同性，决定了随机序列的特性。

为了便于把握，下面给出几条评价密钥流发生器的简单标准。

（1）所产生的密钥序列要足够长。

（2）在密钥流中，0 和 1 的出现频率应接近；如果密钥流是字节流，则 256 种可能字节的出现频率应接近。

（3）种子密钥也应有足够的长度，最少不能小于 128b。

现在人们已经开发出了多种性能良好的密钥流生成器，所采用的方法有以下几种。

① 线性反馈移位寄存器（Linear Feedback Shift Register，LFSR）。

② 非线性移位寄存器（NLFSR）。

③ 有限自动机。

④ 线性同余。

⑤ 混沌密码序列等。

比较有名气的算法是 RSA 数据安全公司的 Ron Rivest 于 1987 年基于移位寄存器方法

设计的 RC4。它是一种同步流密码。

### 2.1.7 量子加密

如前所述，现代加密体系主要有对称加密体系和非对称加密体系两种。在长期的不断发展中，这两个系列都在不断改进、推陈出新。在对称算法中，除了前面介绍的 AES 外，先后还出现过前面提及的美国的 DES 及其各种变种 Triple DES、GDES、New DES 和 DES 的前身 Lucifer，欧洲的 IDEA，日本的 FEALN、LOKI91、Skipjack、RC4、RC5 以及以代换密码和转轮密码为代表的古典密码等。在非对称密码算法中，除了前面介绍的 RSA 外，还出现了背包密码、McEliece 密码、Diffe Hellman、Rabin、OngFiatShamir、零知识证明的算法、椭圆曲线、EIGamal 算法等。在长期的应用中，由于这两种加密体系各自的优缺点，这两类加密算法是被合起来形成混合加密体系使用。

但是，不管是对称加密算法，还是非对称加密算法，它们的共同特点都是依赖现有计算能力不足来保证安全的——实现破解的时间往往要远长于密码保护信息的有效期。这样的条件限制，使它们的安全性随着计算能力的飞速提高日益降低。例如，1975 年由 IBM 公司提出，1976 年被作为数据加密标准的 56 位 DES 加密算法，在当时的运算条件看来，破解是非常困难的，几乎是不可能的。可是以现有运算能力来看，只需要 2 天就能完全破解。再如，美国政府目前应用最广泛的、由美国专门制定密码算法的标准机构——美国国家标准技术研究院与美国国家安全局设计的 SHA-1 密码算法，已经在 2005 年初被中国山东大学王小云教授和她的研究小组宣布成功破解，为此，美国国家标准与技术研究院不得不宣布，美国政府 5 年内将不再使用 SHA-1 密码算法。还有，以现在计算能力认为需要数千年才能破解的 RSA 加密体系，早从 2000 年 8 月就开始走上被破解之路。因为，那时 IBM 完成了一台量子计算机的研发。而唯一能与量子计算机对抗的只有一种技术——量子加密技术。

#### 1. 薛定谔猫和杨氏双孔干涉实验

为了描述量子力学的真相，1935 年奥地利著名物理学家薛定谔提出一个有关猫既是死的又是活的著名思想实验：如图 2.10 所示，在这个实验中，把一只猫放在一个密封的盒中，一种巧妙设计的连锁装置把放射性物质和一个装有剧毒氰化物的玻璃小瓶联系在一起，当放射性样品中的某个原子发生了衰变时，一把榔头就会落下打破瓶子，把盒子里面的猫杀



图 2.10 薛定谔猫

死。现在的情况是，假设下一分钟原子发生衰变的可能性是不确定的，概率各是 50%。现在的问题是，一分钟之后这个猫是死是活？

如果人们开启盒子，就会发现该猫非死即生。但是在此之前，猫的量子态应是死猫状态和活猫状态的混合。也就说处于活猫-死猫量子叠加态中。因此，只要不对这个盒子里的猫做观测，那么这个猫就永远保持不死不活的状态，而对这个盒子中的猫只要看上一下，这个猫的死活就决定了，也就是说猫的不死不活状态就不存在了。这个佯谬的结果就是，对于猫的不死不活状态是不可观测的，观测的行为本身就破坏了这个状态。

这是一个微观的量子态的宏观描述，实际上，在微观尺度上，对量子的观察会造成不可避免的干扰，这就是不确定原理，即粒子的某些特性无法确定，直到测量外力迫使它们选择。或者说，如果没有揭开盖子，进行观察，则永远也不知道猫是死是活，它将永远处于半死不活——既死又活的叠加态。打开盒子，标志着叠加态突然结束（在数学术语就是“波函数坍缩（collapse）”）。

与薛定谔猫极为相似的是 1801 年托马斯•杨进行的一次光学实验——著名的杨氏双孔干涉实验。在这个实验中，光显示出波粒二象性。后来又有人做了单电子双缝干涉实验，在一张纸板上开两道狭缝，然后在纸板后架上光屏，用一个仪器发射出一个电子，使其通过狭缝。按照正常思维来看，电子可能通过这个狭缝，也可能通过另一个狭缝。而观测显示：这个电子在两道狭缝后都形成了图像。有人说只射出一个电子，也会产生干涉条纹，这是德布罗意物质波的相干叠加所导致的必然结果。

对于这些现象如何解释呢？1957 年，埃弗雷特提出的“多世界诠释”为人们打开了一条新的理论通道。埃弗雷特认为两只猫都是真实的。有一只活猫，有一只死猫，但它们位于不同的世界中。格利宾显然十分赞赏这一诠释，他说：“在量子的多世界中，我们通过参与而选择出自己的道路。在我们生活的这个世界上，没有隐变量，上帝不会掷骰子，一切都是真实的。”

### 2. 海森堡测不准原理和单量子不可复制定理

“海森堡测不准原理”是量子力学的基本原理，它表明，在同一时刻以相同的精度测定量子的位置与动量是不可能的，只能精确测定两者之一。因为测量一个粒子的位置和速度，其办法是将光照到这粒子上，一部分光波被此粒子散射开，用于指明它的位置。人们不可能将粒子的位置确定到到光的两个波峰之间距离更小的程度，故必须用短波长的光来测量，至少要用一个光量子。这量子会扰动这粒子，并改变粒子的速度，而且位置测量得越准确所需的波长就越短，单独量子的能量就越大，粒子的速度就被扰动得越厉害，即对粒子的位置测量得越准确，对速度的测量就越不准确。

“单量子不可复制定理”是“海森堡测不准原理”的推论，它表明，在不知道量子状态的情况下复制单个量子是不可能的，因为要复制单个量子就只能先作测量，而测量必然改变量子的状态，所以说不可能。可利用量子的这些特性来解决秘密密钥分发的难题。

### 3. 量子加密的基本原理

量子加密技术是构建在量子不确定性原理上的。其工作原理如下：使用量子加密法的

两个用户各自产生一个私有的随机数字符串。第一个用户向第二个用户的接收装置发送代表数字字符串的单个量子序列（光脉冲），接收装置从两个字符串中取出相匹配的比特值。这些比特值就组成了密钥的基础。

量子加密法的先进之处在于这种方法依赖量子力学定律。传输的光量子是无法进行窃听的，量子要么被接收者的接收机接收，要么被窃听者接收，因为，如果有人进行窃听，窃听动作本身将会对通信系统造成干扰，对通信系统的量子状态造成不可挽回的变化。通信双方就会得知有人进行窃听，从而结束通信，生成新的密钥。

**4. 量子加密技术的发展与现状**

量子加密概念是在 1970 年由美国科学家威斯纳首先提出的。当时的想法是利用单量子态制造不可伪造的“电子钞票”。但这个设想的实现需要长时间保存单量子态，不太现实。因此长时间不受重视。直到 1984 年，第一个量子密码术方案（BB84）被提出。才真正迎来了量子加密技术的新时代。

1992 年，一种新的更简单的方案（B92）被提出，量子密码技术不被设计用于传输密文，而被设计用于传输密钥。根据量子力学的不确定性原理以及量子不可克隆定理，任何窃听者的存在都会被发现，从而保证密钥的绝对安全，也就保证了加密信息的绝对安全。

最初的量子密码通信利用的都是光子的偏振特性，目前主流的实验方案则用光子的相位特性进行编码。目前，在量子密码术实验研究上进展最快的国家为英国、瑞士和美国。英国国防研究部于 1993 年首先在光纤中实现了基于 BB84 方案的相位编码量子密钥分发，光纤传输长度为 10km。

到 1995 年，在 30km 长的光纤传输中成功实现了量子密钥分发。与偏振编码相比，相位编码的好处是对光的偏振态要求不那么苛刻。在长距离的光纤传输中，光的偏振性会退化，造成误码率的增加。

美国洛斯阿拉莫斯国家实验室以 B92 方案成功地在长达 48km 的地下光缆中传送量子密钥，同时他们在自由空间里也获得了成功。

1999 年，瑞典和日本合作，在光纤中成功地进行了 40km 的量子密码通信实验。而在中国，量子密码通信的研究刚刚起步，中科院物理所于 1995 年以 BB84 方案在国内首次做了演示性实验，华东师范大学用 B92 方案做了实验，但也是在距离较短的自由空间里进行的。2000 年，中科院物理所与研究生院合作，在 850nm 的单模光纤中完成了 1.1km 的量子密码通信演示性实验。2002 年德国慕尼黑大学和英国军方下属的研究机构合作，在德国和奥地利边境的楚格峰和卡尔文德尔峰之间用激光成功传输了光子密钥，传输的距离达到 23.4km。从 2003 年开始，位于日内瓦的 id Quantique 公司和位于纽约的 MagiQ 技术公司，推出了传送量子密钥的距离超越了贝内特实验中 30cm 的商业产品。IBM、富士通和东芝等企业也在积极进行研发。市面上的产品能够将密钥通过光纤传送几十千米。

就在世界各国都热衷于量子加密的研究与时间时，瑞典林雪坪大学（Linkoping University）信息编码学系教授 Jan-Ake Larsson 与博士生 Jonathan Jogenfors 在研究能量时间缠结技术时发现，如果使用传统光源取代光子，那么窃听者就能找到密钥——编码字串。这就使窃听者能顺利读取资料，而不会被侦测到。这说明已经成为许多量子加密系统基础

的“能量-时间缠结”方法并非牢不可破，也存在着易于攻击的“漏洞”。随后斯德哥尔摩大学（Stockholm University）的物理学家也通过实验，证实光源完全可能被更换掉，造成数据可能被窃听。不过，人们也乐观地认为这个问题是可以解决的。

## 2.1.8 信息隐藏

### 1. 数据隐藏及其处理过程

信息隐藏（information hiding）是指隐蔽数据的存在性，通常是把一个秘密信息（secret message）隐藏在另一个可以公开的信息载体（cover）之中，形成新的隐秘载体（stego cover）。目的是不让非法者知道隐秘载体中是否隐藏了秘密信息，并且即使知道也难以从中提取或去除秘密信息。

信息隐藏与数据加密都是用来保护信息机密性的手段，并且信息隐藏技术也承袭了数据加密的一些基本思想和概念，例如信息隐藏的过程也可以利用密钥进行控制。但是，信息隐藏与数据加密所采用的技术手段不同。数据加密的基本方法是编码，通过编码将明文变换为密文。而信息隐藏是“隐藏”，使非法者难以找到秘密信息。一般多用多媒体数据作为载体。这是因为多媒体数据本身具有极大的冗余性，具有较大的掩蔽效应。

图 2.11 表明了信息的隐藏过程和提取过程。

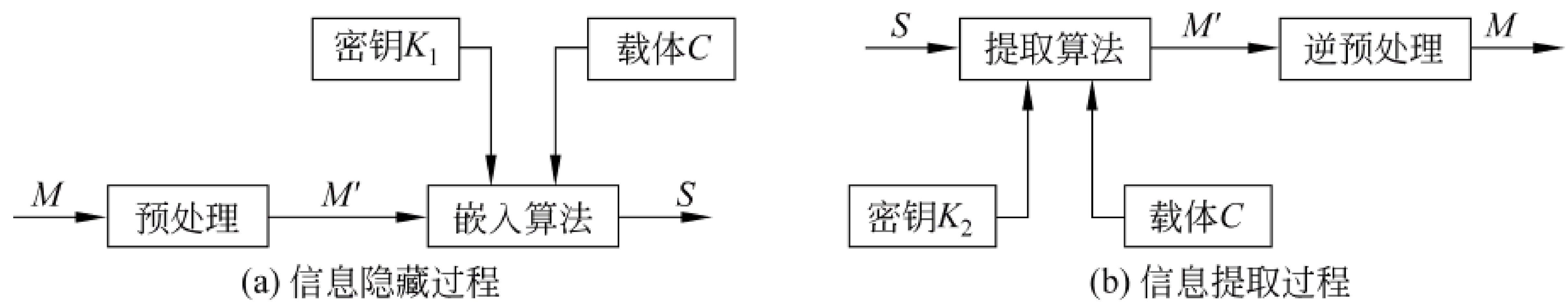



图 2.11　信息的隐藏过程和提取过程

信息隐藏过程如下。

（1）对原始报文 $M$ 进行预处理（如加密、压缩等）形成隐藏报文 $M'$。

（2）在密钥 $K_1$ 的控制下，通过嵌入算法（embedding algorithm）将隐藏报文 $M'$ 隐藏于公开信息载体 $C$ 中，形成隐秘载体 $S$。

信息提取过程如下。

（1）在密钥 $K_2$ 的控制下，使用提取算法从隐秘载体 $S$ 中提取出隐藏报文 $M'$。

（2）对隐藏报文 $M'$ 进行解密、解压等逆预处理，恢复出原来的报文 $M$。

### 2. 信息隐藏技术分类

对信息隐藏技术可以进行如下分类。

1）按照载体类型分类

按照载体类型可将信息隐藏技术分为如下几类。

（1）文本载体信息隐藏。

（2）图像载体信息隐藏。

（3）声音载体信息隐藏。

（4）视频载体信息隐藏。
（5）二进制流载体隐藏等。

2）按照控制密钥分类

按照控制密钥可将信息隐藏技术分为如下几类。
（1）对称隐藏算法。
（2）公钥隐藏算法。

3）按照隐藏位置分类

按照隐藏位置可将信息隐藏技术分为如下几类。
（1）信道隐藏。利用信道固有的特性进行信息隐藏。目前主要有两类：基于网络模型的信息隐藏和基于扩频的信息隐藏。
（2）空域/时域信息隐藏。利用待隐藏信息位替换载体中的一些最不重要的位。例如，把表示一个像素点灰度的数值中的 180 替换为 181，不会产生太大影响。
（3）变换域信息隐藏。把待隐藏信息嵌入到载体的变换空间（如频域）中。这种方法具有分布性、可变换性和较高的鲁棒性。

## 2.2 消息认证——完整性保护

### 2.2.1 数据完整性保护与消息认证

#### 1. 数据完整性保护的概念

数据的完整性保护是针对如下 3 种攻击所采取的措施。
（1）内容篡改（content modification），包括对报文内容的插入、删除、改变等。
（2）序列篡改（sequence modification），包括对报文序列的插入、删除、错序等。
（3）时间篡改（timing modification），对报文进行延迟或回放。
也就是说，数据的完整性包括了内容完整性、序列完整性和时间完整性。

#### 2. 消息认证的概念

消息认证（message authentication）也称为报文鉴别，是用于验证所收到的消息确实来自真正的发送方，并未被篡改，检测传输和存储的消息（报文）有无受到完整性攻击的手段，包括消息内容认证、消息的序列认证和操作时间认证等。其核心是消息（报文）的内容认证。消息的序列认证的一般办法是给发送的报文加一个序列号，接收方通过检查序列号来鉴别报文传送的序列有没有被破坏。消息的操作时间认证也称为数据的实时性保护，通常可以采用时间戳或询问-应答机制进行确认。

从功能上看，一个消息认证系统分为两个层次：低层是认证函数，上层是认证协议。关于认证协议在后面讨论，这里主要讨论认证函数。认证函数的功能是能够由报文产生一个鉴别码。

**3. 鉴别码的传递与用法**

图 2.12 给出鉴别码的 4 种传送与使用方法。图中的 $F$ 为鉴别码生成函数，$F(M)$为鉴别码，$K$ 为对称加密密钥，SK 和 PK 为非对称加密的私钥和公钥，$E$ 为加密，$D$ 为解密。

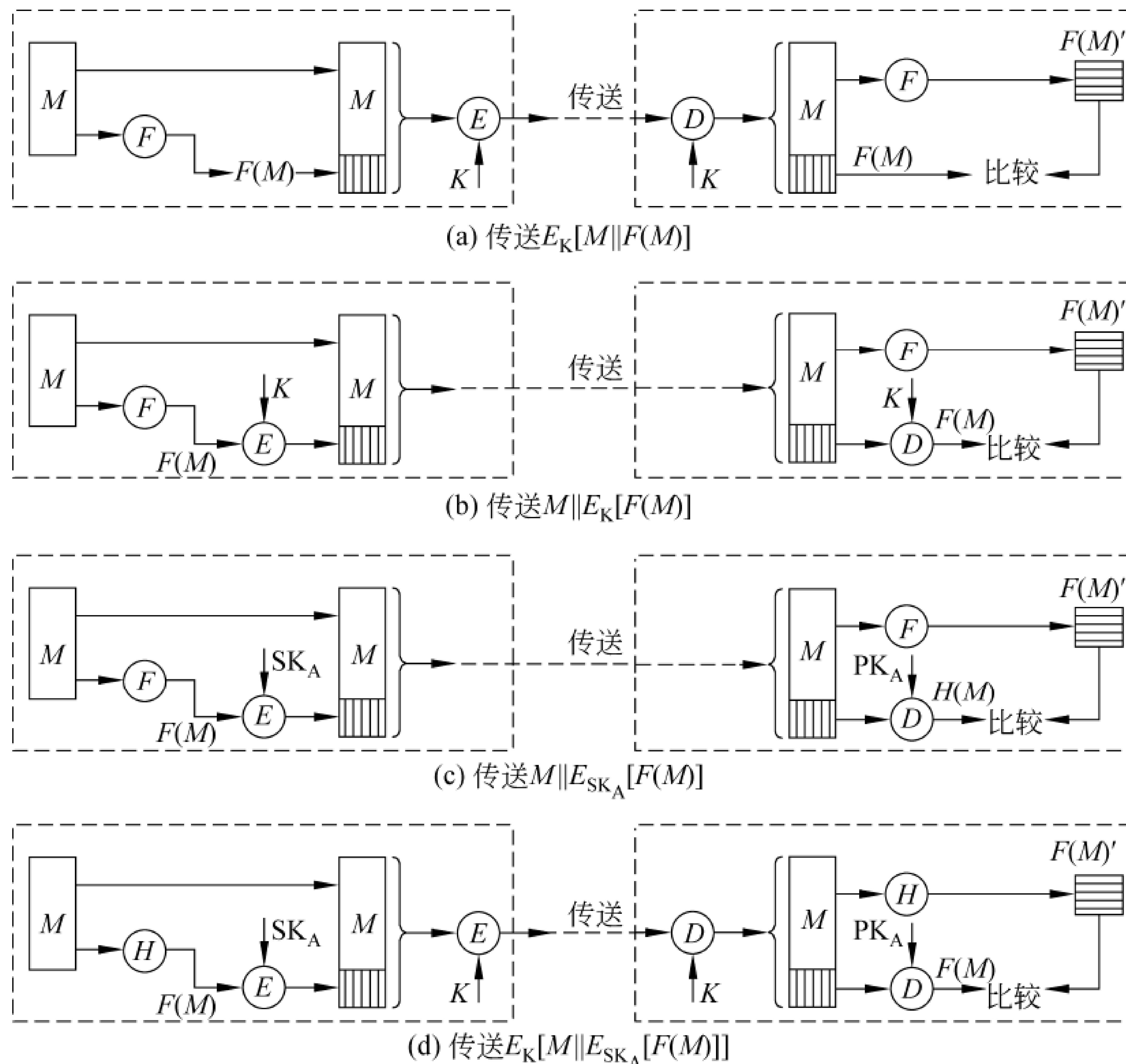



图 2.12　鉴别码的 4 种传送与使用方法

可以看出，使用鉴别码可以在接收方进行报文是否受到内容完整性攻击的鉴别依据。其中，图 2.12（a）和图 2.12（d）用于对报文有机密性保护的场合，图 2.12（b）和图 2.12（c）用于对报文不要求机密性保护的场合。当然还可以有其他传送与使用方法，这里不再介绍。

**4. 鉴别码生成函数**

进行完整性保护的基本思路是从原来的报文生成一个具有唯一性的鉴别码，并且传递是秘密的。这样将报文及其鉴别码一同传送到目的方后，用同样的方法从接收的报文再生成一次鉴别码，由于鉴别码具有唯一性，比较两个鉴别码，就能证实报文在传送过程中有没有被删除、插入或修改过。

按照这一思路，若将报文的密文与明文一同传送，接收方再从明文生成一次密文，再对两个密文进行比较，就可以进行报文的完整性鉴别。但是，这种鉴别码太长，传送效率太低。

传统的冗余校验码也可以看作是一种简单的消息认证方法。但它主要用于防止人工操作或传输中的偶然错误。若用于对付攻击者，就勉为其难了。因为这些算法比较简单而且

是公开的，技术熟练者完全可以成功地绕开这些检测。

目前广为采用的生成报文鉴别码的方法主要有两种。

（1）基于消息认证码（Message Authentication Code，MAC）的方法。MAC 也称为密码校验和，是使用一个由密钥控制的鉴别码生成函数基于报文产生的固定长的鉴别码。图 2.13 为基于 MAC 的消息认证过程，其中 *C* 为 MAC 生成函数。

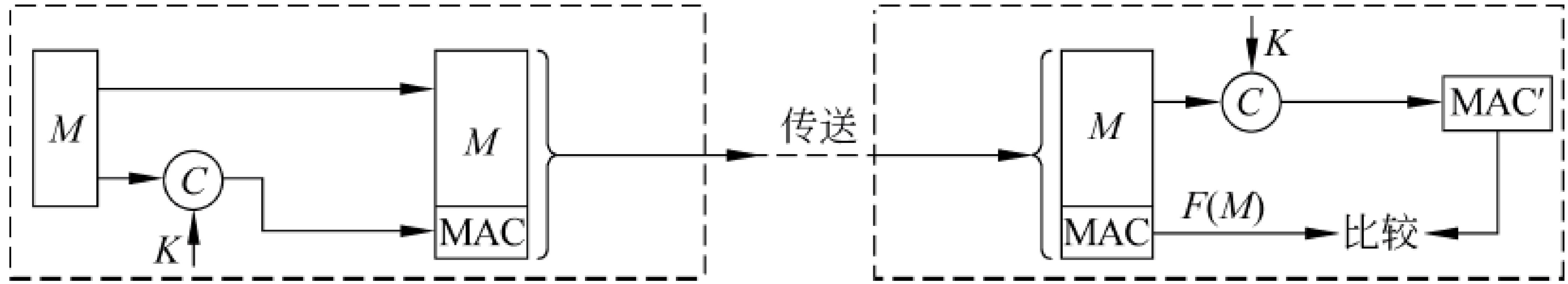



图 2.13　基于 MAC 的消息认证过程

（2）基于报文（消息）摘要（Message Digest，MD）的方法。报文摘要也称为单向杂凑码或单向散列码，是将报文使用单向杂凑（hash，也译为哈希或散列）函数变换成为具有固定长度的鉴别码。它具有两个主要特点：一是不使用密钥，仅仅是输入消息的函数；二是具有错误检测能力，输入报文中任何一位或多位的改变都会导致其摘要（散列码）的改变。图 2.14 为将一个报文 *M* 利用杂凑函数 *H* 得到 MD 的过程。

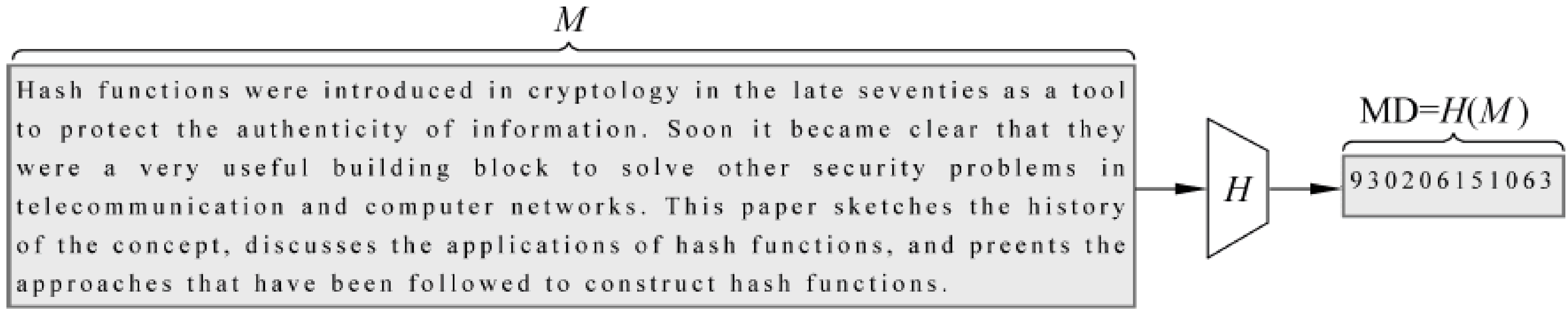



图 2.14　报文摘要的生成

## 2.2.2　MAC 函数

### 1．MAC 函数的特点

MAC 函数是消息 *M* 和密钥 *K* 的函数，即 $\text{MAC} = C(M, K)$ 或 $\text{MAC} = C_K(M)$。从形式上看，MAC 函数与对称加密函数极为类似，都需要一个信源端和信宿端共享的密钥。但是，它们又有本质上的区别。

（1）加密算法要求可逆性，而 MAC 算法不要求可逆性。

（2）加密函数明文长度与密文长度一般相同，是一对一的函数，而 MAC 函数则是多对一的函数，其定义域由任意长的消息组成，而值域则由 MAC 比特的比特位构成。若报文长度为 $m$ 位，MAC 长度为 $n$ 位，则有 $2^m$ 种报文对应 $2^n$ 种 MAC。由于 $m >> n$，所以一定存在不同的报文产生相同的 MAC。例如，使用 100b 的报文和 10b 的 MAC，则有 $2^{100}$ 种报文对应 $2^{10}$ 种 MAC，平均而言，$2^{100}/2^{10} = 2^{90}$ 个报文具有相同的 MAC。

（3）MAC 函数比加密函数更不容易攻破，因为即便攻破，也无法验证其正确性。

### 2. 安全的 MAC 函数应具备的性质

一个安全的 MAC 函数应具备下列性质。

（1）对于已知的 $M$ 和 MAC，要找到满足 MAC′ = MAC 的另一个 $M'$，在计算上是不可能的。

（2）$C_K(M)$应当是均匀分布的，则对于任何随机选择的报文 $M$ 和 $M'$，找到 $C_K(M') = C_K(M)$ 的概率是 $2^{-n}$，$n$ 为 MAC 的位数。

（3）若 $M'$是 $M$ 的一个已知变换，则找到 $C_K(M') = C_K(M)$ 的概率是 $2^{-n}$。

因此，若只有收发双方才知道密钥 $K$，并且接收到的 MAC 与计算出的 MAC 相等，则接收方可以相信：

（1）接收到的消息未被修改。

（2）接收到的消息来自真正的发送方。

（3）如果消息中含有序列号，则消息的顺序也是正确的。

**3. 安全的 MAC 函数对密钥长度的要求**

安全的 MAC 函数要求密钥具有足够的长度。这样可以使得攻击者用穷举方法确定 MAC 密钥成为不很容易的事情。

从 MAC 认证的理论上看，仅需要一个源、宿共享的密钥。但这种情况下只能提供认证而不能提供保密性保护，因为报文是以明文传送的。若将 MAC 附在报文后面对整个报文进行加密，则既可提供认证又可提供保密。但这需要两个独立的密钥，并要被源、宿两方共享。

**4．CBC-MAC**

CBC-MAC 也称为数据认证算法，是建立在 DES 基础上，基于分组密码，并按照密文块链接（Cipher Block Chaining，CBC）模式操作的 MAC 构造方法之一，也是 ANSI 的一个标准，如图 2.15 所示。

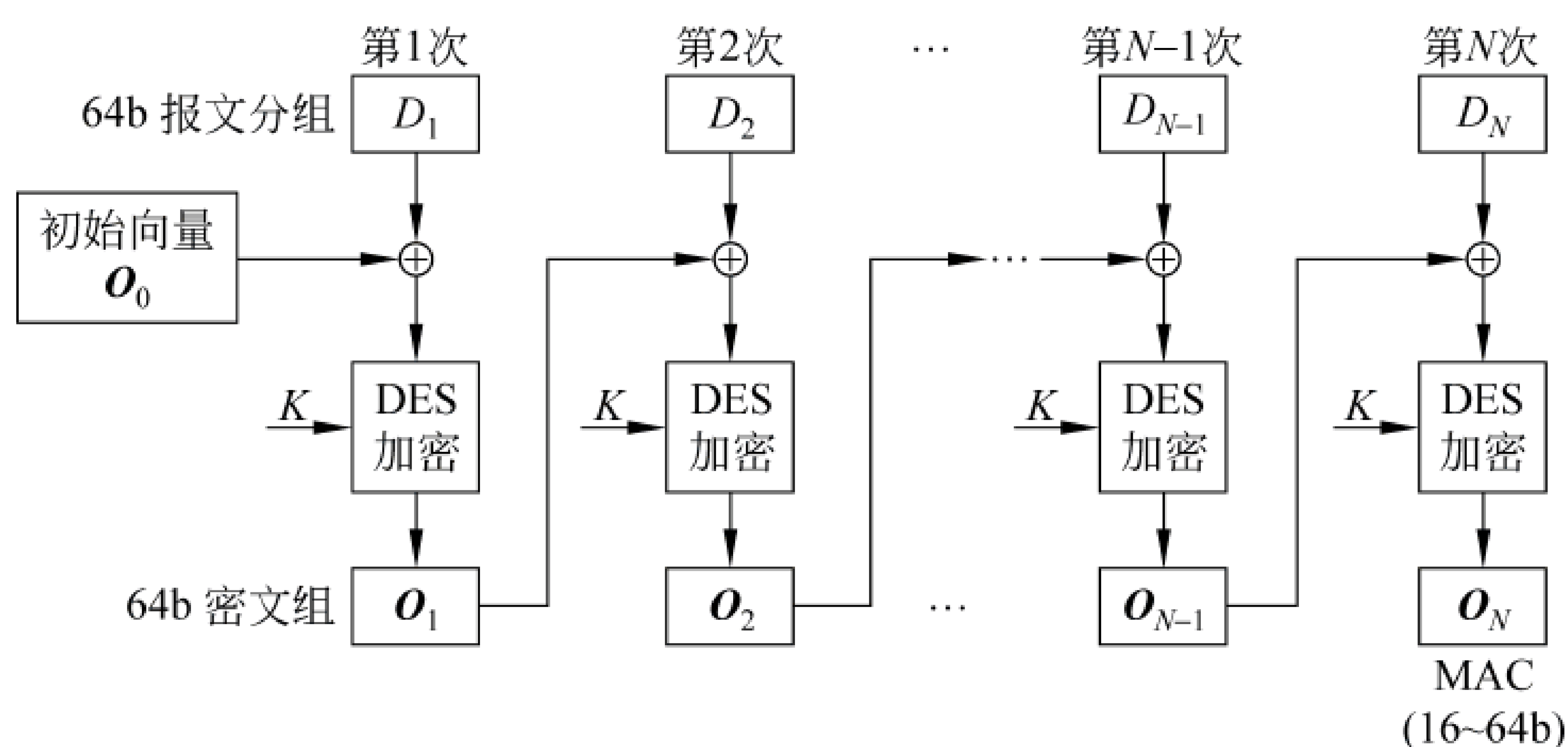



图 2.15　数据认证算法

CBC 模式的基本特点如下。

（1）将要认证的数据分成连续的 64b 的分组 $D_1$、$D_2$、…、$D_N$，若最后的分组不足 64b，在其后加 0，补足 64b。

（2）每个分组在用密钥加密之前要先与前一组的密文组进行异或运算生成其加密过程的种子向量。初始向量 $\boldsymbol{O}_0$ 取零。

（3）最后的 MAC 用 $\boldsymbol{O}_N$ 最左边 $M$ 位表示，并且 $16 \leqslant M \leqslant 64$。

### 2.2.3　哈希函数

哈希（Hash）函数也称为散列函数或杂凑函数，它能把任意长的输入消息串（又称为预映射，pre-image）通过哈希算法变换成固定长度的输出串，该输出就是哈希值。简单地说，哈希函数就是一种将任意长度的消息压缩到某一固定长度的消息摘要的函数。

#### 1. 对哈希函数的一般要求

报文摘要就像是需要鉴别的数据的一个“指纹”。为了实现对于消息的鉴别，哈希函数应具备以下性质。

（1）$H()$可应用于任意长度的输入数据块，产生固定长度的哈希值。

（2）对于每一个给定输入数据 $M$，计算出它的哈希值 $h = H(M)$很容易。

（3）给定哈希值 $h$，倒推出输入数据 $M$ 在计算上不可行，即单向性。

（4）对于给定的报文 $M$ 和其哈希值 $h$，要找到另一个 $M' \neq M$，使 $H(M')=H(M)$极其困难，即弱抗碰撞（collision）性。

（5）找到任何满足 $H(M)=H(M')$且 $M \neq M'$的报文对$(M, M')$在计算上是不可行的，即强抗碰撞性。

碰撞性是指对于两个不同的报文，如果它们的摘要值相同，则发生了碰撞。

性质（1）是将哈希函数应用于消息认证的实际需求。

性质（2）和（3）是单向性，由给定消息产生哈希值很简单，而给定哈希值则不可能产生对应的消息。

性质（4）保证无法找到一个替代报文，使它的哈希值与给定消息产生的哈希值相同，能防止伪造。

性质（5）指的是哈希函数对已知的生日攻击方法的防御能力。

哈希函数并不提供机密性，并且它们不使用密钥以生成摘要。哈希函数非常适合于认证和确保数据的完整性。

**例 2.7**　设报文有 $m$ 组分组，哈希码 $C$ 的长度为 $n$ 位，则某一位哈希码 $C_i$ 可以这样简单地计算：

$$C_i = B_{i1} \oplus B_{i2} \oplus \cdots \oplus B_{im}$$

当然，它并不完全满足对哈希函数的要求。典型的报文摘要算法有 MD5（Riverst 提出，1992 年 RFC 1321 公布，码长 128b）和安全散列算法 SHA（Secure Hash Algorithm，码长 160b）。由于 SHA 比 MD5 多了 32b，所以更安全，但要慢些。

#### 2. 报文摘要算法 MD5

MD5 是沿着 MD2、MD4 的轨迹进化形成的报文摘要算法的最新版本。它的开发者是作为 RSA 算法设计者之一的 Ronald L.Riverst。MD5 不以任何假设和密码体制为基础，是一个直接构造出来的算法。它的主要特点如下。

（1）单向性由报文生成报文摘要，但不能由报文摘要还原为报文。

（2）无碰撞性。对于不同的报文不会产生两个相同的报文摘要。

（3）运算速度快，应用比较普遍。

1）MD5 的主要应用

MD5 主要应用在如下两个方面：防篡改鉴别和加密。

MD5 的典型应用是对一段报文产生一个 128b 的报文摘要。例如，在 UNIX 下有很多软件在下载的时候都有一个扩展名为 md5 的文件，在这个文件中通常只有一行文本，大致结构如下：

```
MD5 （tanajiya.tar.gz） = 0ca175b9c0f726a831d895e269332461
```

这就是 tanajiya.tar.gz 文件的数字签名。如果在以后传播这个文件的过程中，无论文件的内容发生了任何形式的改变（包括人为修改或者下载过程中线路不稳定引起的传输错误等），只要对这个文件重新计算 MD5 值就会发现报文摘要不相同，由此可以确定得到的只是一个不正确的文件。如果再有一个第三方的鉴别机构，用 MD5 还可以防止文件作者的“抵赖”，这就是数字签名的应用。

在加密系统中，密码的保管非常重要。由于具有系统管理员权限的用户可以读密码文件，所以常规保存的密码文件对具有系统管理员权限的用户就无秘密可言。而 MD5 的单向性和无碰撞性使得用户密码可以用 MD5（或其他类似的算法）加密后存储。当用户登录的时候，系统把用户输入的密码计算成 MD5 值，然后再去和保存在文件系统中的 MD5 值进行比较，进而确定输入的密码是否正确。这样，系统就可以在不知道用户密码的明码的情况下确定用户登录系统的合法性，不但可以避免用户的密码被知道，而且还在一定程度上增加了密码被破解的难度。

2）MD5 算法描述

MD5 算法的基本轮廓如下。

（1）以 512b 分组来处理输入的报文。

（2）每一分组又被划分为 16 个 32b 子分组。

（3）经过了一系列的处理后，输出 4 个 32b 分组放在 4 个链接变量（chaining variable）或称寄存器 $A$、$B$、$C$、$D$ 中。

（4）将 4 个链接变量进行级联，生成一个 128b 的哈希值。

上述轮廓可以分为如下 4 步完成。

① 数据扩展。将报文按照图 2.16 的格式用 100…0 进行填充，并附加一个 64b 的原始报文长度字段，使得总长度为 512b 的整数倍。应当注意，填充是必需的，若报文长度为（512b 的整数倍–64b），则还需填充一个 512b 长度的填充字段。

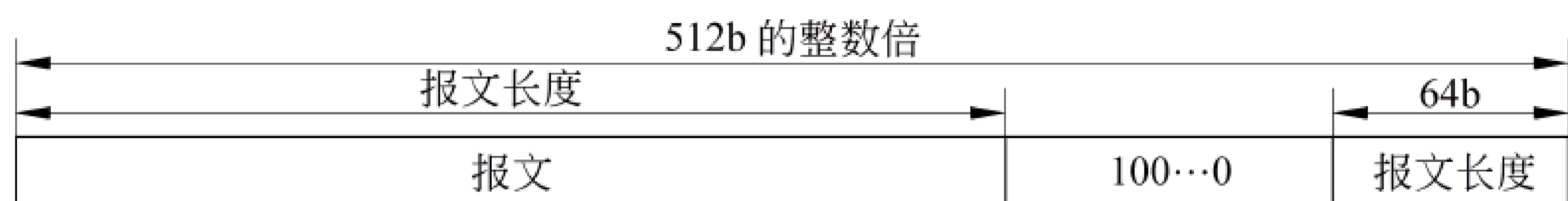



图 2.16　数据准备格式

② 初始化 MD 缓冲区。使它们的十六进制初始值分别为

$A = 0x01234567$，$B = 0x89abcdef$，$C = 0xfedcba98$，$D = 0x76543210$

③ 报文切块。按照 512b 的长度，将报文分割成 $N+1$ 个分组：$Y_0$、$Y_1$、…、$Y_N$。每一分组又可以表示为 16 个 32b 的字。

④ 依次对各分组进行 $H_{\mathrm{MD5}}$ 压缩生成 MD5 值。如图 2.17 所示，从分组 $Y_0$ 开始到最后一个分组 $Y_N$，依次进行 $H_{\mathrm{MD5}}$ 压缩运算。每个 $H_{\mathrm{MD5}}$ 有两个输入（一个 128b 的 $\mathrm{CV}_q$ 和一个 512b 的分组 $Y_q$）和一个 128b 输出；最开始的 128b 输入为 4 个 32b 的链接变量；以后每个 $H_{\mathrm{MD5}}$ 的输出作为下一个 128b 输入。最后一个 128b 输出即所求的 MD5 值。显然，这个过程可以用循环或递归实现。

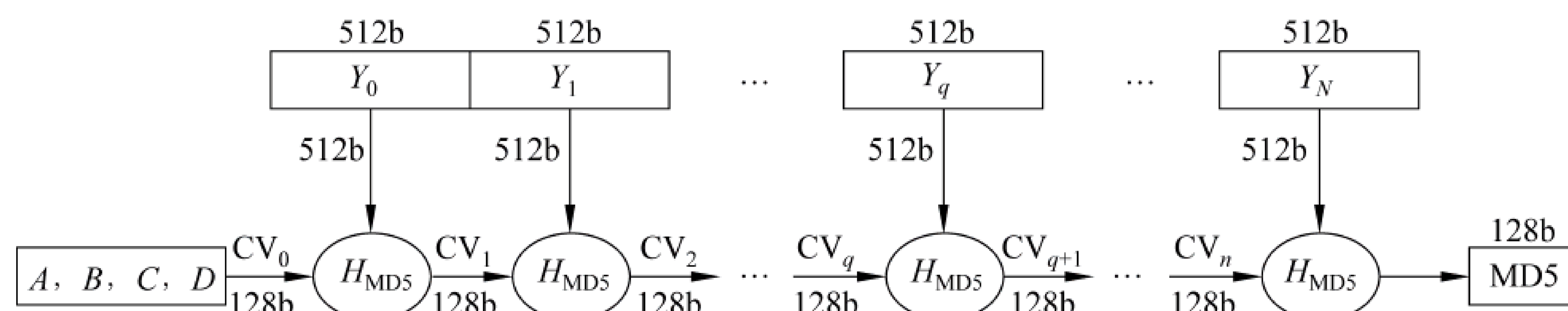



图 2.17　依次对各分组进行 $H_{\mathrm{MD5}}$ 压缩生成 MD5 值

$H_{\mathrm{MD5}}$ 算法是 MD5 的关键函数。它的计算由 4 轮处理组成，每一轮又包括了 16 个操作，总共 64 次操作，每一次操作要使用一个常数。关于它的细节这里不再介绍。需要说明的是，MD5 曾经被人们认为是无碰撞的。但是，在 2004 年 8 月 17 日在美国加州圣巴巴拉召开的国际密码学会议（Crypto’2004）上，我国山东大学教授王小云宣布她已经找到了使 MD5 产生碰撞的方法。但是，目前人们还没有找到 MD5 的合适的代替方案。

### 3. 安全哈希算法

安全哈希算法（Secure Hash Algorithm，SHA）是由 NIST 和 NSA 开发的，于 1993 年作为美国联邦信息处理标准（FIPS PUB 180）公布，1995 年修订为 FIPS PUB 181，1995 年修订为 SHA-1。另外还有 4 种变体曾经发布，以提升输出的范围和变更一些细微设计：SHA-224、SHA-256、SHA-384 和 SHA-512（名称中的数字代表摘要长度）。

SHA 基于 MD4 算法，SHA 的设计很近似于 MD4 模型。SHA 在用于数字签名的标准算法（DSS）中也是安全性很高的一个哈希算法。该算法的输入为小于 264b 长的任意消息，分为 512b 长的分组，输出为 160b 长的消息摘要。因为它能产生 160b 的哈希值，所以，抗穷举攻击能力更强。

SHA 与 MD5 都是来自 MD4，所以它们有许多相似之处。这里对 SHA 的细节不进行介绍，仅在表 2.7 中对这两种报文摘要算法进行比较。

**表 2.7　MD5 与 SHA-1 的比较**

| 算　法 | 摘要长度/b | 最大报文长度 | 分组处理长度/b | 运算次数 | 常数个数 |
|---|---|---|---|---|---|
| MD5 | 128 | 无限制 | 512 | 64（4 轮 16 次） | 64 |
| SHA-1 | 160 | $2^{64}-1$ | 512 | 80（4 轮 20 次） | 4 |

总之，SHA-1 比 MD5 抗击穷举搜索的强度高，但执行速度较慢。

### 实验 6　实现报文认证算法

**1. 实验目的**

（1）加深对报文认证算法的理解（如哈希函数的概念及密钥指纹的生成方法）。

（2）掌握报文认证算法的应用方法。

**2. 实验内容**

（1）运行一种报文认证算法（MD5、SHA 等）程序。

（2）测试运行的报文认证算法程序。

（3）分别按照下面的一种模式使用上述报文认证算法并进行比较。

① 报文与鉴别码链接后，用单钥加密传送。

② 仅对鉴别码用单钥加密，然后再与被认证报文链接传送。

③ 用公钥加密算法使用发方私钥仅对鉴别码加密，然后再与被认证报文链接传送。

④ 用公钥加密算法使用发方私钥仅对鉴别码加密，与被认证报文链接，再用单钥加密后传送。

⑤ 将报文与通信双方共享的秘密值 S 连接后计算鉴别码，附加到被认证报文上发送。

**3. 实验准备**

（1）可根据不同要求，选择进行下面的工作之一。

① 下载一种报文认证算法的程序源代码，进行解析。

② 自己设计实现一种报文认证程序。

（2）编译、调试上述程序。

（3）设计完成实验内容的环境和步骤。

**4. 推荐的分析讨论内容**

（1）你知道哪些报文认证算法？试进行比较。

（2）分别分析报文认证算法的几种不同使用模式对于保证数据在机密性、完整性以及在防范伪造、抵赖、冒充、篡改风险方面的作用。

（3）这种认证算法安全吗？有没有看到关于它的可攻击的报道？试到网上搜索一下。

（4）其他发现或想到的问题。

## 2.3　数 字 签 名

### 2.3.1　数字签名及其特征

在日常生活中，为了确认一件作品及其源出处，常需要采取签名、落款、骑缝章等手段，以便于鉴别。在数据通信过程中，有时会发生一方对另一方的如下一些欺骗行为。

（1）否认。发送方否认自己发送过的某个报文，或接收方接收一个报文后，否定接收过。

（2）冒充。发送方冒充第三方给接收方发送报文。

（3）伪造。某一方自己伪造一份报文，却声称来自对方。

（4）篡改。接收方收到一份报文后进行修改，却说这是对方发来的报文原样。

面对这些问题，在通信双方尚未建立起信任关系且存在利害冲突的情况下，单纯的报文鉴别是无能为力的。为此不得不采用数字签名（digital signature）。

在 ISO 7498-2 标准中，数字签名定义为："附加在数据单元上的一些数据，或是对数据单元所做的密码变换，这种数据和变换允许数据单元的接收者用于确认数据单元来源和数据单元的完整性，并保护数据，防止被人（例如接收者）进行伪造。"

数字签名是实现数据的不可否认性保护的机制。为了达到这一目的，它应当具有与手工签名一样的性质。

（1）签名是可信的。

（2）签名是无法被伪造的。

（3）签名不可以重复使用。

（4）签名以后不可以被篡改。

（5）签名具有不可否认性。

从有效性和可行性出发，对数字签名技术有以下要求。

（1）签名的结果必须是与签名的报文相关的二进制位串。

（2）签名要能够验证签名者的身份以及签名的日期和时间。

（3）签名能够用于证实被签报文的内容的真实性。

（4）签名的产生、识别和验证应比较容易。

（5）数字签名应当可以被备份。

（6）用已知的签名构造一个新的报文或由已知的报文产生一个假冒的签名，在计算上都是不可行的。

（7）签名可以由第三方验证，以解决双方在通信中的争议。

### 2.3.2 直接数字签名

所谓直接方式，就是签名过程只有发送方和接收方参与。实施这种方法的前提是接收方可以通过某种方式验证发送方提交的凭证，也可以在发生争议时将该凭证交第三方仲裁。

那么，用什么作为直接数字签名的凭证呢？

好像对称密钥就可以作为数字签名中的凭证。可以设想，如果发送方 A 和接收方 B 使用只有双方才使用的密钥，那么 A 向 B 发送了一份加密的报文，B 就可以断定是 A 发来的。因为除了 A 和 B，没有第三方知道这个密钥。但是，由于 A 和 B 都知道这个密钥，所以只能抵御冒充，无法抵御否认、篡改和伪造。

与之相比，非对称密钥就要好得多。采用非对称密钥，接收方 B 需要知道发送方 A 的私钥。这样，A 用自己的私钥将报文加密，将其传送到 B 后，B 用 A 的公钥将密文解密。这个过程中，A 不可能冒充第三方，也无法否认自己的发送；B 也无法篡改和伪造。

但是，非对称密钥算法的效率是很低的，不宜用于长的报文的加密。为此可以采用鉴别码，将报文 M 通过一个单向哈希函数生成短的定长鉴别码，将认证与签名结合起来进行。

直接签名后的报文有可能被接收方滥用。例如，A 发送给 B 一张电子支票，有可能被 B 多次复制兑换现金。如果 A 在报文中再增加一种特有凭证，如时间戳（timestamp），就可以避免这种情况发生。

另外，直接签名方法将一种算法既用于认证又用于签名，使签名的有效性完全依赖于密钥体制的安全性，也会形成一些漏洞。例如，发送方会声称自己的密钥被窃或被盗用，来否认已经发送报文。为避免这样的威胁，可以要求每一个被签名的报文都要包含一个时间戳，标明报文发送的日期和时间，同时要求一旦密钥丢失或被盗用，要立即向管理机构报告并更换密钥。但是，若密钥真正被盗，盗窃者可以伪造一个报文，并加上一个他盗窃密钥之前的时间戳。

### 2.3.3 有仲裁的数字签名

有仲裁的数字签名的基本思想是：发送方完成签字后，不是直接发送给接收方，而是将报文和签字先发送给双方共同信任的第三方进行验证，第三方验证无误后，再附加一个“已经通过验证”的说明并注上日期，一同发送给接收方。由于第三方的介入，发送方和接收方都无法抵赖。

有仲裁的数字签名方法也有很多，下面仅举几例。假定报文由 X 向 Y 传送，A 为仲裁者，$M$ 为传输报文，$H(M)$为哈希函数值，||为链接，$\mathrm{ID_X}$ 为 X 的身份码，$T$ 是时间戳，则可以有如下几种方案。

**1. 方案 1**

（1）X→A：$M \| E_{K_{\mathrm{XA}}}[\mathrm{ID_X} \| H(M)]$。X 将签名 $E_{K_{\mathrm{XA}}}[\mathrm{ID_X} \| H(M)]$和报文 $M$ 发给 A。$K_{\mathrm{XA}}$ 为 X 和 A 的共享密钥。

（2）A→Y：$E_{K_{\mathrm{AY}}}[\mathrm{ID_X} \| M \| E_{K_{\mathrm{XA}}}[\mathrm{ID_X} \| H(M)] \| T]$。A 对签字验证后，再附加上时间戳 $T$ 并用 A 和 Y 共享的密钥 $K_{\mathrm{AY}}$ 加密后转发给 Y。

（3）Y 收到 A 发来的报文，解密后，将结果保存起来。由于 Y 不知道 $K_{\mathrm{XA}}$，所以不能直接检查 X 的签字，只能相信 A。

当出现争议时，Y 可以声称自己收到的 $M$ 来自 X，并将

$$E_{K_{\mathrm{AY}}}[[\mathrm{ID_X} \| M \| E_{K_{\mathrm{XA}}}[\mathrm{ID_X} \| H(M)] \| T]$$

发送给 A，让 A 仲裁。

这个方案是建立在 X 和 Y 都对 A 高度信任的基础上。

（1）X 相信 A 不会泄露 $K_{\mathrm{XA}}$，也不会伪造自己的签名。

（2）Y 相信 A 所验证 X 的签字是可靠的。

（3）X 和 Y 都相信出现争议时 A 能公正地处理。

所以会得到如下结论。

（1）X 相信 Y 无法对收到的报文予以否认。

（2）Y 相信 X 不会对他所发送的报文予以否认。

但是，这个方案未提供保密性，X 传送给 A 的 $M$ 是明文形式。此外，方案本身没有对

仲裁者的限制机制，一旦仲裁者不公正，如与发送方共谋否认发送过的报文，或与接收方连手伪造发送方的签字，都会形成签字的漏洞。

**2. 方案 2**

（1）X→A：$ID_X \| E_{K_{XY}}[M] \| E_{K_{XA}}[ID_X \| H(E_{K_{XY}}[M])]$。

（2）A→Y：$E_{K_{AY}}[ID_X \| E_{K_{XY}}[M] \| E_{K_{XA}}[ID_X \| H(E_{K_{XY}}[M])] \| T]$。

这个方案用 X 和 Y 的共享密钥 $K_{XY}$ 加密所传送的 $M$，从而提供了保密性。但还没有解决对仲裁者的约束。

**3. 方案 3**

（1）X→A：$ID_X \| E_{SK_X}[ID_X \| E_{PK_Y}[E_{SK_X}[M]]]$。

（2）A→Y：$E_{SK_A}[ID_X \| E_{PK_Y}[E_{SK_X}[M]] \| T]$。

在这个方案中，仲裁者只能用 X 的公钥对 $E_{SK_X}[ID_X \| E_{PK_Y}[E_{SK_X}[M]]]$ 解密，得到 $ID_X'$ 与以明文形式传送来的 $ID_X$ 进行比较，确认这个报文确实来自 X，却不能解密 $E_{PK_Y}[E_{SK_X}[M]]$。$E_{PK_Y}[E_{SK_X}[M]]$ 要由 Y 才能解密。因为 Y 可以使用 $PK_A$ 解密 $E_{SK_A}[ID_X \| E_{PK_Y}[E_{SK_X}[M]] \| T]$，进一步用 $SK_Y$ 解密 $E_{PK_Y}[E_{SK_X}[M]] \| T]$，再用 $PK_X$ 解密 $E_{SK_X}[M]$。这样就使仲裁方无法与任何一方共谋。

### 2.3.4 数字签名标准 DSA

数字签名算法（Digital Signature Algorithm，DSA）是美国国家标准委员会（NIST）公布的数字签名标准（Digital Signature Standard，DSS）。DSA 最早公布于 1991 年，在征求了公众意见后在 1993 年和 1996 年又发布了两次修改版。图 2.18 描述了 DSA 签名的基本过程。

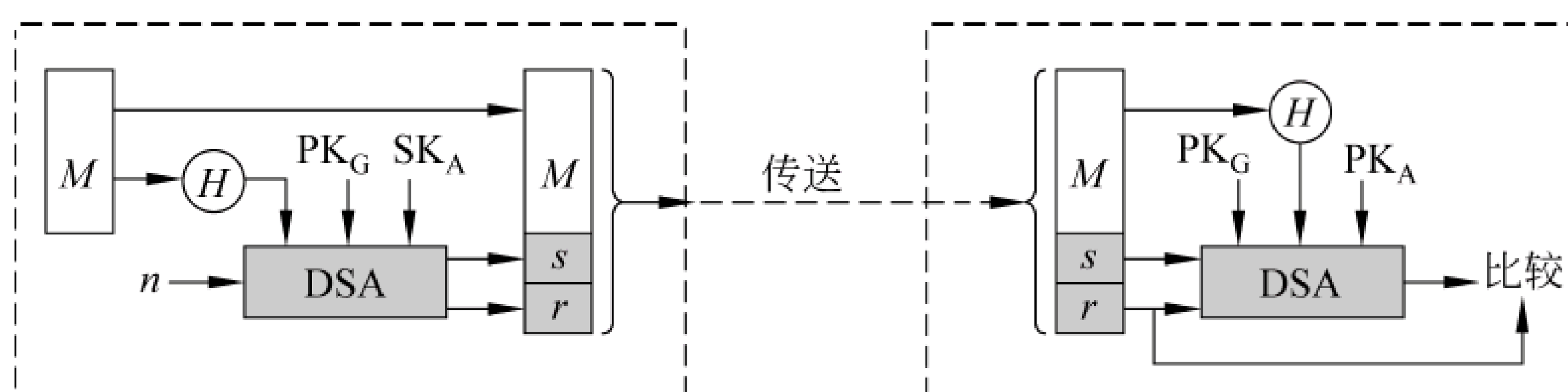



图 2.18 DSA 签名的基本过程

DSA 算法的签名函数以以下参数作为输入。

（1）用 SHA 方法生成的报文摘要。

（2）一个随机数。

（3）发送方的私有密钥 $SK_A$。

（4）全局公钥 $PK_G$——供所有用户使用的一族参数。

DSA 算法输出两个数据：$s$ 和 $r$。这两个输出就构成了对报文 $M$ 的数字签名。

接收方收到报文后，先产生出报文的摘要，再将这个摘要和收到的签名以及全局公钥

$PK_G$、发送方的公开密钥 $PK_A$ 一起送到 DSA 的验证函数中，生成一个新的 $r'$。若 $r'$与 $r$ 相等，就说明签字有效。

DSA 算法的安全性不再依赖于加密密钥的安全性。同时，其计算基于求离散对数的困难性，使攻击者从 $r$ 恢复 $n$，或从 $s$ 恢复 $SK_A$ 都是在计算上不可行的。所以，DSA 比采用 RSA 的签名方法要可靠得多。

除基于离散对数的签名算法外，人们还开发了其他一些签名算法。关于它们的细节，这里不再介绍。

### 2.3.5 认证协议实例——SET

安全电子交易（Secure Electronic Transaction，SET）协议是一种利用加密技术（cryptography），以确保信用卡消费者、销售商及金融机构在 Internet 上从事电子交易的安全性和隐私性的协议。它是由两大信用卡公司 VISA 和 Master 在 GTE、IBM、Microsoft、Netscape、SAIC、Terisa System、Verisign 等著名 IT 公司的支持下开发的，于 1996 年 2 月 1 日正式发布。

#### 1. 电子支付中的安全需求

支付是交易的重要环节。由于电子支付的虚拟性，它的安全性备受人们关注，也是电子商务安全最重要和最困难的环节。归纳起来，电子支付主要有如下一些安全需求。

（1）确定交易过程中交易伙伴的真实性。

（2）保证电子单据的隐秘性，防范被无关者窃取。

（3）保证业务单据不丢失，即使丢失也可察觉。

（4）验证电子单据内容的完整性。

（5）验证电子单据内容的真实性。

（6）具有防抵赖性和可仲裁性。

（7）保证存储的交易信息的安全性。

#### 2. SET 的目标

SET 的主要目标如下。

（1）保证数据在 Internet 上安全传输，不被窃取。

（2）保证数据的完整性，即保证数据在传输过程中不被篡改。

（3）订单信息与账号信息相隔离，在将包括消费者账号信息的订单送给商家时，商家应只看到订货信息，而看不到消费者账号信息。

（4）参与各方可以互相认证，不仅消费者与商家可以互相认证（一般由第三方提供信用担保），消费者、商家和银行之间也能够相互认证。

（5）采用统一的协议和数据格式，使不同厂家开发的软件具有兼容性和互操作性，并可以运行在不同的硬件和操作系统平台上。

### 3．SET 的参与角色

一个 SET 交易过程可能会涉及如下 6 种角色。

（1）持卡者（cardholder），即消费者，必须具备如下条件。

① 持有发卡机构支付卡账号。

② 持有认证机构颁发的身份证书。

③ 能够上网，可以使用支付卡结算。

（2）商家（merchant），即经营者，是商品或服务的提供者，必须具备如下条件。

① 拥有网上经营许可权。

② 持有银行委托授权机构（认证中心、CA）颁发的数字证书。

③ 具有的电子商务平台能处理消费者申请，与支付网关通信，存储自身公钥签名，存储自己的公钥交换私钥，存储交易参与方的公钥交换私钥，申请和接受认证，与后台数据库通信等。

（3）银行（acquirer）：给在线交易参与者建立账号，并处理支付卡的认证和支付业务。

（4）发卡机构（issuer）：为每一个建立了账户的客户颁发支付卡，也可以由指派的第三方承担。

（5）支付网关（payment gateway）：将 Internet 上的传输数据转换为金融机构内部数据，并实现商家和持卡人的身份验证。

支付网关必须具有如下条件。

① 银行授权。

② CA 颁发的数字证书。

（6）SET 认证中心：为了使交易参加方有一个可信的第三方，建立了一个认证中心（CA）来为消费者、商家和银行颁发数字证书，分配密钥。

SET 认证中心采用层次结构。其最高层 CA（即 Root CA）的安全维系着整个 SET 协议的安全。为了防止 Root CA 遭破解，SET 将其密钥长度定为 2048b，比一般用户的密钥长度 1024b 长一倍。

### 4．SET 的安全保障作用

（1）基于持卡人的安全保障包括以下几方面。

① 对账号数据保密。

② 对交易数据保密。

③ 持卡人对商家的认证。

（2）基于商家的安全保障包括以下几方面。

① 对交易数据完整性的认证。

② 对持卡人账户数据的认证。

③ 对持卡人的认证。

④ 对银行端的认证。

⑤ 对交易数据保密。

⑥ 提供交易数据的佐证。

（3）基于银行（支付网关）的安全保障包括以下几个方面。

① 对消费者账号数据的认证。

② 对交易数据的间接认证。

③ 对交易数据完整性的认证。

④ 提供交易数据的佐证。

### 5. 电子商务中的 B to C 交易过程

交易指买卖双方之间进行商品交易的行为。广义的交易是一个复杂的过程。在电子商务中，由于参与方不同等原因，会形成不同的交易过程，如 B to B（Business to Business，企业间电子商务，也缩写为 B2B）的交易过程、B to C（Business to Consumer，企业对消费者的电子商务，也缩写为 B2C）的交易过程等。下面主要介绍 B to C 的交易过程。

图 2.19 为 B to C 的基本交易过程。它可以被分为如下 4 种业务 7 个步骤。

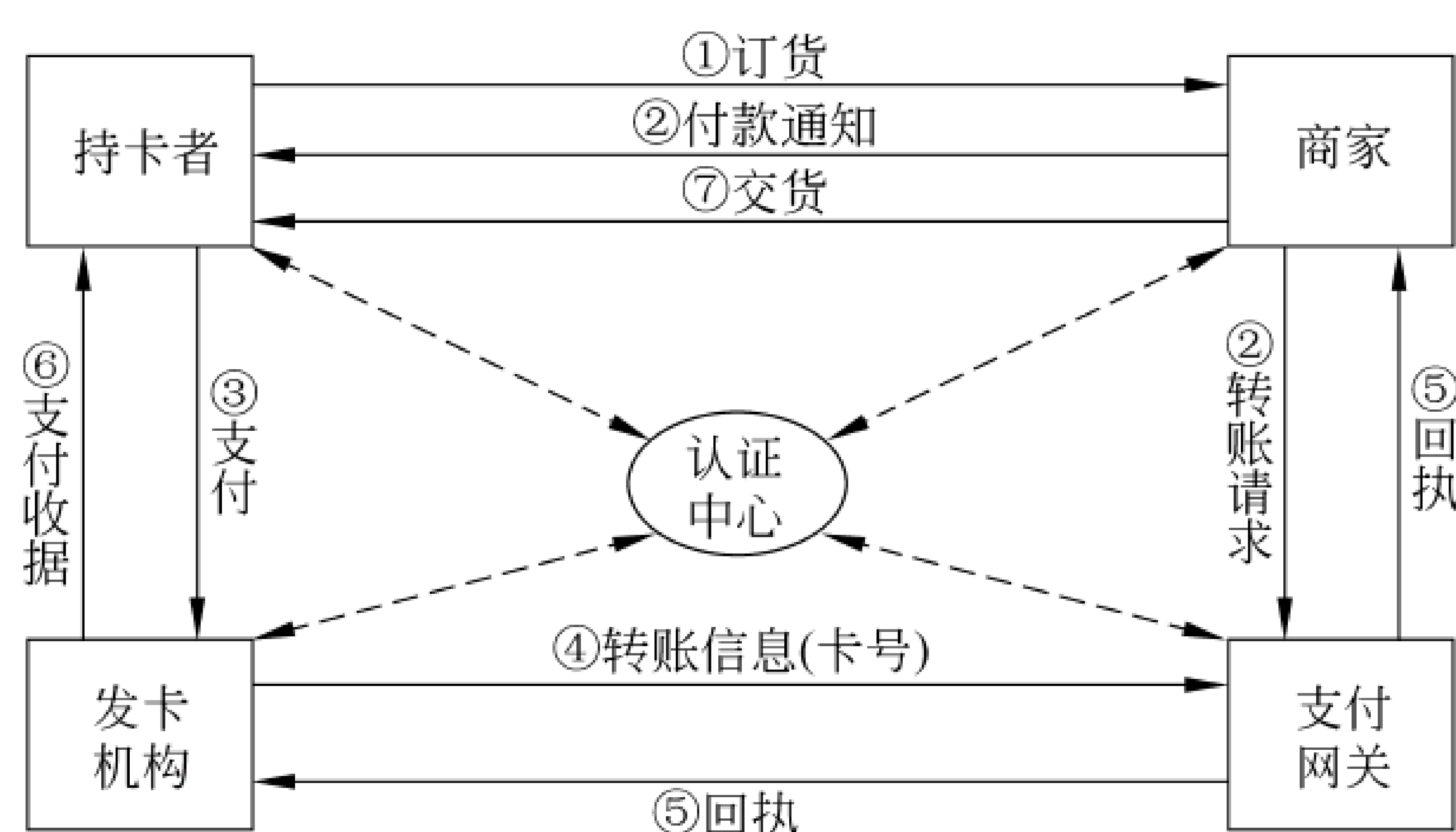



图 2.19　B to C 的基本交易过程

（1）订货。

①：消费者上网，查看企业和商家网页，选择商品；消费者通过对话框填写订货单（姓名、地址、品种、规格、数量和价格）并提交给商家。

（2）支付。

②：商家核对消费者订货单后，向用户发出付款通知，同时向银行发出转账请求。

③：消费者选择支付方式，如向发卡机构提交支付卡。

（3）转账。

④：发卡机构验证消费者的支付卡后，将卡号加密传向银行（支付网关）。银行审查消费者支付卡（有效、款额等）合格后，进行转账。

⑤：转账成功，向商家和发卡机构发出转账成功回执。

⑥：发卡机构向消费者发出支付收据。

（4）付货。

⑦：商家向消费者交货。

### 6．SET 关键技术之一——数字信封

数字信封是为了解决密钥传送的安全而产生的技术。图 2.20 是用数字信封传递对称密钥的过程。这个对称密钥由发送方随机产生。




图 2.20　用数字信封传递对称密钥的过程

实际上，这个信封的制作非常简单，就是用接收方的公钥对要传送的信息（这里是对称密钥）进行加密。打开信封的方法是用接收方的私钥进行解密。

### 7．SET 关键技术之二——双重签名

SET 有一个基本性能：不需要让与有关角色无关的其他角色知道的机密，就不让它知道，例如：

① 只与持卡人和商家之间的交易有关的机密数据（如订单信息 OI），不必要也不可以让银行知道。

② 持卡人的账户数据也是持卡人的个人秘密数据（如付款指示 PI），不必要也不可以让商家知道。

但是，在这种情况下，还要让商家和银行都可以确定这些数据确实是由持卡人产生和送出的，可以间接或直接地对这些数据进行认证。这一性能在 SET 中使用双重签名（dual signature）技巧实现。双重签名的基本过程如图 2.21 所示，具体实现时略有所不同。下面介绍两种方案。

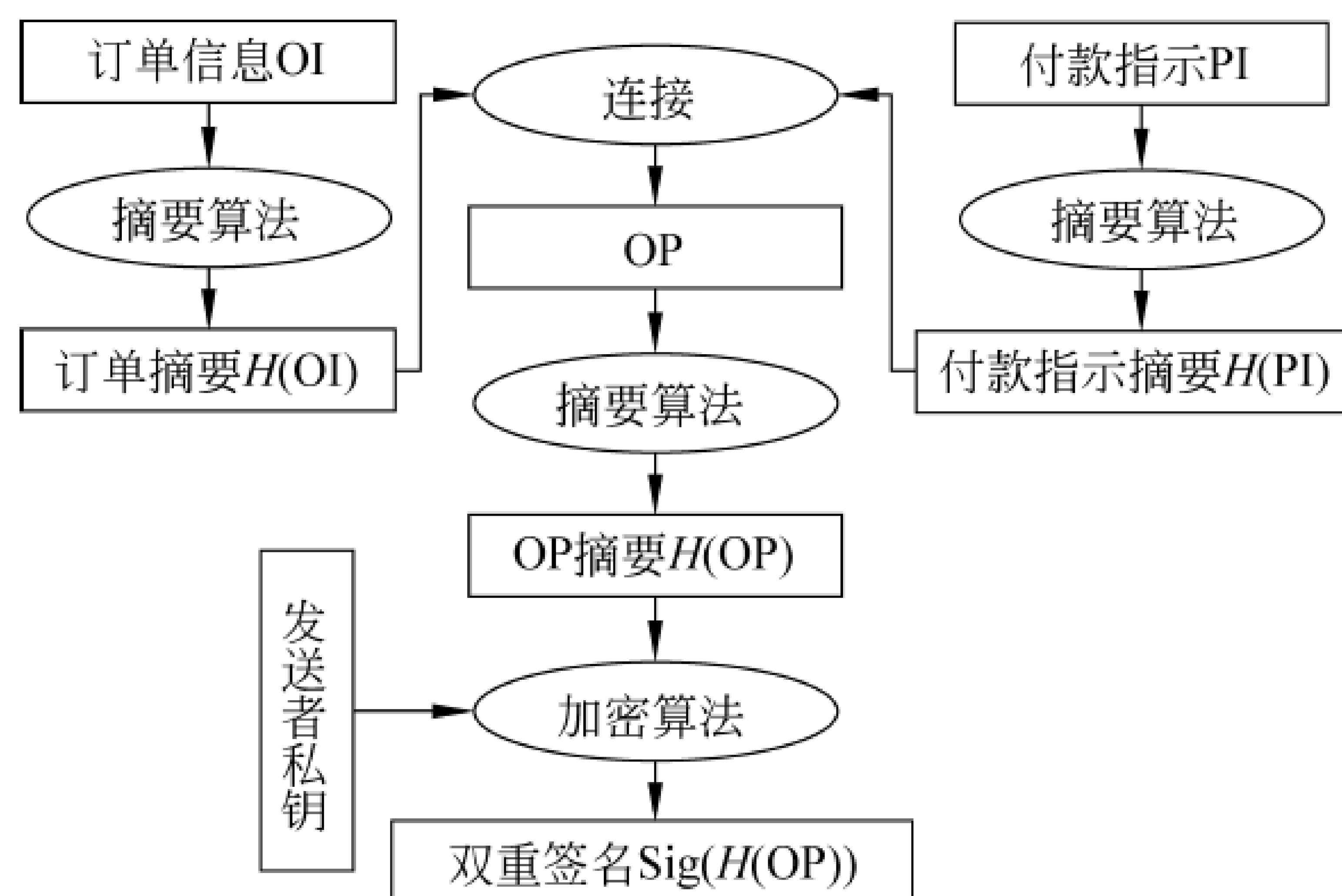



图 2.21　持卡人进行的双重签名过程

1）方案 1

设 Sig($x$)是一个基于 RSA 的多项式时间函数，可以产生对 $x$ 的有效签名。

（1）持卡人（C）的操作如下。

① 产生订货信息 OI 和账号信息 PI，生成两个摘要 $H_C$(OI)和 $H_C$(PI)。

② 双重签名：将 $H_C$(OI)和 $H_C$(PI)连接，形成 OP(OI+PI)，再生成一个摘要 $H_C(H_C(OI) \| H_C(PI))$，并对其用私钥进行签名得到 Csig($H_C(H_C(OI) \| H_C(PI))$)。

③ 发给商家：OI、$H_C$(PI)和 Csig($H_C(H_C(OI) \| H_C(PI))$)。

④ 发给支付网关：PI、$H_C$(OI)和 Csig($H_C(H_C(OI) \| H_C(PI))$)。

（2）商家（M）收到信息后的操作如下。

① 利用收到的 OI，生成摘要 $H_M$(OI)。

② 将 $H_M$(OI)与 $H_C$(PI)合起来生成摘要 $H_M(H_M(OI) \| H_C(PI))$。

③ 用 $H_M(H_M(OI), H_C(PI))$对 Csig($H_C(H_C(OI) \| H_C(PI))$)进行验证，以确认信息发送者的身份和信息是否被修改过。

（3）支付网关（P）收到信息后的操作如下。

① 利用收到的 PI，生成摘要 $H_P$(PI)。

② 将 $H_C$(OI)与 $H_P$(PI)合起来生成摘要 $H_P(H_C(OI) \| H_P(PI))$。

③ 用 $H_P(H_C(OI), H_P(PI))$对 Csig($H_C(H_C(OI) \| H_C(PI))$)进行验证，以确认信息发送者的身份和信息是否被修改过。

2）方案 2

（1）持卡人（C）的操作如下。

① 生成 OI、PI、$H$(OI)、$H$(PI)、Csig($H$(OI))、Csig($H(H(OI) \| H(PI))$)。

② 将 Csig($H$(OI))、Csig($H(H(OI) \| H(PI))$)、$H$(PI)和 OI 传送给商家。

③ 将 PI 和 $H$(OI)传给支付网关（银行）。

（2）商家（M）操作如下。

① 用 C 的公钥验证 Csig($H(H(OI) \| H(PI))$)、Csig($H$(OI))和 $H$(PI)，确定是否持卡者的签名。

② 用自己的私钥生成对交易数据的签名 Msig($H$(OI))。

③ 将 Msig($H$(OI))和 Csig($H(H(OI) \| H(PI))$)一同送支付网关（银行）。

（3）支付网关（P）操作。

① 验证 Msig($H$(OI))，确定 $H$(OI)为交易数据的哈希值。

② 用已知的 PI 和 $H$(OI)，验证 Csig($H(H(OI) \| \mathrm{H}(PI))$)的正确性，确认交易数据和账户数据都是正确的。

③ 根据政策，确认此笔交易是否成功。

图 2.22 为 SET 交易过程。它用双重签名解决了三方参加电子贸易中的安全通信问题。

## 实验 7　加密软件 PGP 的使用

### 1. 实验说明

PGP（Pretty Good Privacy）是一个基于 RSA 公钥的邮件加密软件，也是一个与 Linux 齐名的优秀自由软件。其最早的版本是由美国的 Philip R. Zimmermann 开发的，并于 1991 年在 Internet 上免费发布。为了打破美国政府对于软件出口的限制，PGP 的国际版在美国境

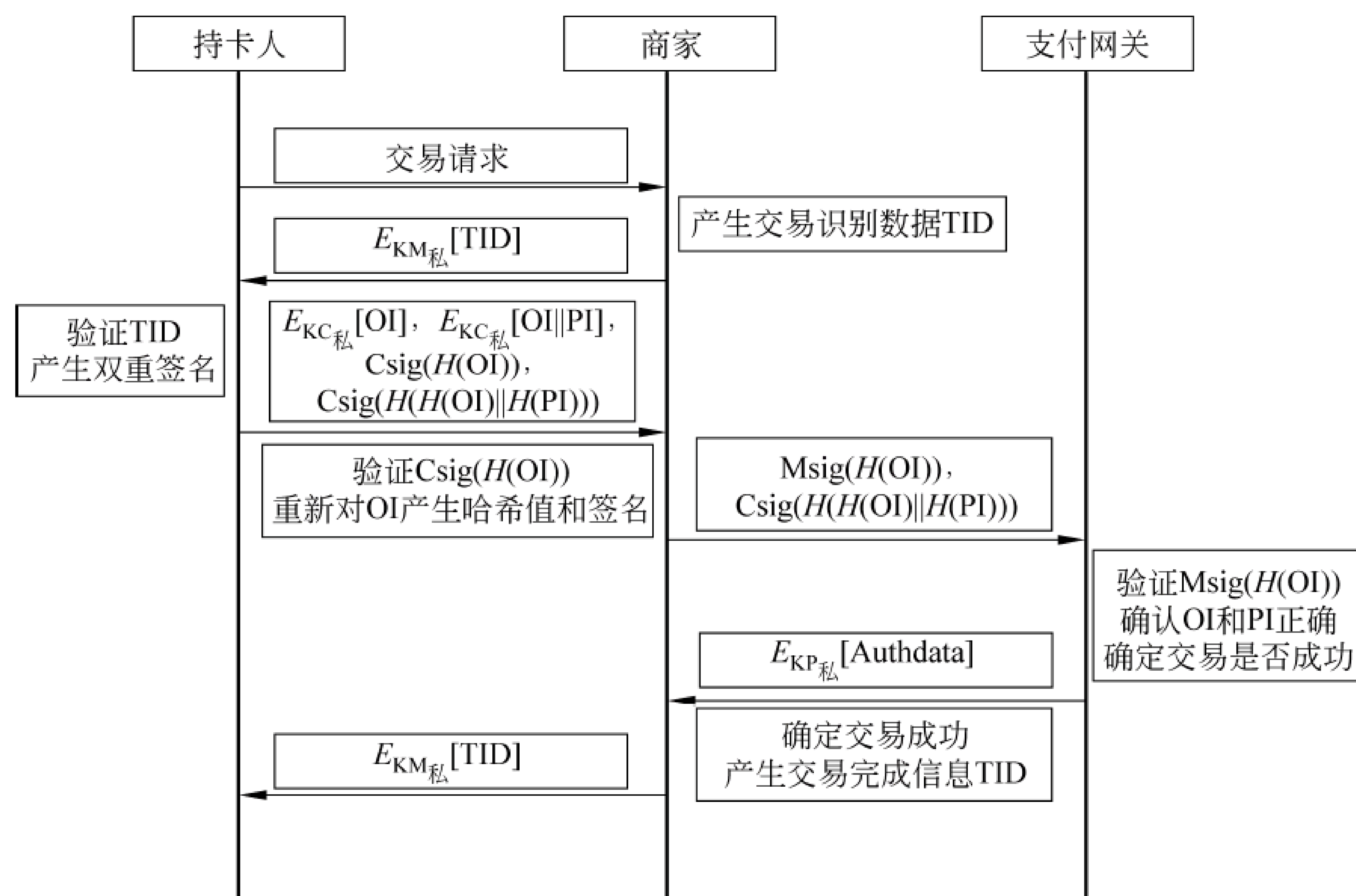



图 2.22　SET 交易过程

注：Authdata 为授权数据。

外开发，并带一个 i 以区别。任何人都可以从挪威的网站 www.pgpi.com 上下载到最新的版本，大小约为 7.8MB。

PGP 采用了审慎的密钥管理，是一种 RSA 和传统加密的杂合算法，用于数字签名的邮件文摘算法以及加密前压缩等，还有一个良好的人机工程设计。它可以让任何人安全地和他从未见过的人们通信，并且事先不需要任何保密的渠道用来传递密钥。同时它的功能强大，有很快的速度，源代码还是免费的。

PGP 对每次会话的报文进行加密后传输，它采用的加密算法包括 AES-256、AES-192、AES-128、CAST、3DES、IDEA 和 Twofish 等。例如，使用 AES 密钥最长可达 256b，这已经足够安全了。

当发送者用 PGP 加密一段明文时，PGP 首先压缩明文，然后建立一个一次性会话密钥，采用传统的对称加密算法（例如 AES 等）加密刚才压缩后的明文，产生密文。然后用接收者的公开密钥加密刚才的一次性会话密钥，随同密文一同传输给接收方。接收方首先用私有密钥解密，获得一次性会话密钥，最后用这个密钥解密密文。

目前，PGP 使用的哈希函数包括 SHA-2（256b）、SHA-2（384b）、SHA-2（512b）、SHA-1（160b）、RIPEMD（128b）、MD5（128b）等。

PGP 在签名之后、加密之前对报文进行压缩。它使用了由 Jean-loup Gailly、Mark Adler、Richard Wales 等编写的 ZIP 压缩算法。

在 PGP 里面最有特色的或许就是它的密钥管理。PGP 包含 4 种密钥：一次性会话密钥、公开密钥、私有密钥和基于口令短语的常规密钥。

用户使用 PGP 时，应该首先生成一个公开密钥/私有密钥对。其中公开密钥可以公开，而私有密钥绝对不能公开。PGP 将公开密钥和私有密钥用两个文件存储，一个用来存储该用户的公开/私有密钥，称为私有密钥环；另一个用来存储其他用户的公开密钥，称为公开

密钥环。

为了确保只有该用户可以访问私有密钥环，PGP 采用了比较简洁和有效的算法。当用户使用 RSA 生成一个新的公开/私有密钥对时，输入一个口令短语，然后使用哈希算法（例如 SHA-1）生成该口令的哈希编码，将其作为密钥，采用 CAST-128 等常规加密算法对私有密钥加密，存储在私有密钥环中。当用户访问私有密钥时，必须提供相应的口令短语，然后 PGP 根据口令短语获得哈希编码，将其作为密钥，对加密的私有密钥解密。通过这种方式，就保证了系统的安全性依赖于口令的安全性。

### 2. 实验目的

（1）掌握 PGP 的下载、安装和使用方法。
（2）进一步加深对于数据加密和数据认证理论的理解。

### 3. 实验内容

（1）下载并安装 PGP。
（2）使用 PGP 生成并管理密钥。
（3）使用 PGP 对文件进行加密/解密。
（4）使用 PGP 对文件进行签名和认证。
（5）使用 PGP 销毁加密文件。

### 4. 实验准备

（1）准备一个实验用的数据文件。
（2）事先浏览可以下载 PGP 软件的网站，了解可以下载的最新 PGP 版本及其特点。
（3）写出实验的详细步骤。
（4）确定下载软件需要的运行环境。

### 5. 推荐的分析讨论内容

（1）你认为 PGP 安全的最大威胁在什么地方？
（2）加密文件的销毁要注意什么？
（3）其他发现或想到的问题。

## 习　　题

### 一、选择题

1. 假设使用一种加密算法，它的加密方法很简单：将每一个字母加5，即a加密成f。这种算法的密钥就是5，那么它属于________。

A. 对称加密技术　　B. 分组密码技术
C. 公钥加密技术　　D. 单向函数密码技术

2. “公开密钥密码体制”的含义是________。

A. 将所有密钥公开　　B. 将私有密钥公开，公开密钥保密

C. 将公开密钥公开，私有密钥保密　　D. 两个密钥相同

3. A方有一对密钥($K_{A_{公}}$ ,$K_{A_{秘}}$)，B方有一对密钥($K_{B_{公}}$ ,$K_{B_{秘}}$)，A方向B方发送数字签名$M$，对信息 $M$ 加密为$M' = K_{B_{公}}(K_{A_{秘}}(M))$。B方收到密文的解密方案是________。

A. $K_{A_{公}}(K_{A_{秘}}(M'))$　　B. $K_{A_{公}}(K_{A_{公}}(M'))$

C. $K_{A_{公}}(K_{B_{秘}}(M'))$　　D. $K_{B_{秘}}(K_{A_{秘}}(M'))$

4. 使用数字签名技术，在发送端是采用________对要发送的信息进行数字签名。

A. 发送者的公钥　　B. 发送者的私钥

C. 接收者的公钥　　D. 接收者的私钥

5. 使用数字签名技术，在接收端采用________进行签名验证。

A. 发送者的公钥　　B. 发送者的私钥

C. 接收者的公钥　　D. 接收者的私钥

6. 数字签名要预先使用单向哈希函数进行处理的原因是________。

A. 多一道加密工序使密文更难破译

B. 提高密文的计算速度

C. 缩小签名密文的长度，加快数字签名和验证签名的运算速度

D. 保证密文能正确还原成明文

7. 设哈希函数 $H$ 有128个可能的输出（即输出长度为128b），如果 $H$ 的 $k$ 个随机输入中至少有两个产生相同输出的概率大于0.5，则 $k$ 约等于________。

A. $2^{128}$　　B. $2^{64}$　　C. $2^{32}$　　D. $2^{256}$

## 二、填空题

1. 密码系统包括以下4个方面：________、________、________和________。

2. ________是加密算法$E$的逆运算。

3. 如果加密密钥和解密密钥相同，这种密码体制称为________体制。

4. ________算法的安全是基于分解两个大素数的积的困难。

5. 公开密钥加密算法的用途主要包括两个方面：________和________。

6. 密钥管理的主要内容包括密钥的________、________、________、________、________、________和________。

7. 密钥生成形式有两种：一种是由________，另一种是由________。

8. 密钥的分配是指产生并使________获得密钥的过程。

9. 密钥分配中心的英文缩写是________。

10. 消息认证是验证________，即验证数据在传送和存储过程中是否被篡改、重放或延迟等。

11. ________是笔迹签名的模拟，是一种包括防止源点或终点否认的认证技术。

12. ________是实现交易安全的核心技术之一，它的实现基础就是加密技术，能够实现电子文档的辨认和验证。

13. MAC函数类似于加密，它与加密的区别是其________。

14. ________是可接受变长数据输入，并生成定长数据输出的函数。

## 三、问答题

1. 有明文can you understand。

（1）假定有一个密钥，其顺序为2，4，3，1的列换位密码，其换位密文是什么？

（2）设密钥是 $i = 1, 2, 3, 4$ 的一个置换 $f(i) = 1, 3, 4, 2$，则周期为4的换位密文是什么？

2. 设 $P$ = blue sky and green tree，密钥 $K$=data，则采用维吉尼亚密码的加密字母是什么？ASCII编码输出为什么？

3. 比较两种密钥体制的优缺点。

4. 解释AES解密算法。

5. 编写程序，实现AES加密算法。

6. 具有$N$个节点的网络如果使用公开密钥密码算法，每个节点的密钥有多少？网络中的密钥共有多少？

7. 在非对称密码体制中，第三方如何断定通信者有无抵赖或伪造行为？

8. 设通信双方使用RSA加密体制，接收方的公开密钥是$(e, n)$=(5, 35)，求明文 $M = 30$对应的密文。

9. 在使用RSA公钥的通信中，若截取了发送给其他用户的密文 $C$ =10，并且用户的公钥为$(e, n)$= (5, 35)，求对应的明文。

10. 什么是序列密码和分组密码？

11. 简述通信双方如何使用密钥体制建立通信中的信任关系。

12. 有哪些建立公开密钥体制的方法？

13. 常规加密密钥的分配有几种方案？请对比它们的优缺点。

14. 如何利用公开密钥加密进行单钥加密密钥的分配？

15. 什么是无碰撞单向哈希函数？

16. 单钥加密体制的密钥分配有哪几种方法？分别有什么优缺点？

17. 假定两个用户A、B分别与密钥分配中心KDC（Key Distribution Center）有一个共享的主密钥$K_A$和$K_B$，A希望与B建立一个共享的一次性会话密钥，需要完成哪些步骤？

18. 什么是密钥的分层控制？这种方法有什么优点？

19. 无中心的密钥分配时，两个用户A和B建立会话密钥需经过哪几步？

20. 单钥体制中的密钥控制技术是什么？

21. 公钥管理机构向用户A、B分配公钥共有哪几步？

22. 请自己设计一个密钥生成算法，并验证其密钥空间的安全性。

23. 在密钥的生存期间内，如何对密钥进行有效的管理？

24. 销毁被撤销的密钥时应注意些什么？

25. KDC在密钥分配过程中充当何种角色？

26. 简述信息隐藏的基本嵌入和检测过程。

27. 简述数字水印的定义和内容。

28. 简述数字隐藏技术中隐含的信任关系。

29. 收集国内外有关加密或信息隐藏技术的最新动态。

30. 分析消息认证码可能遭受的攻击。

31. 数字签名有什么作用？

32. 描述报文鉴别码和哈希码的区别。

33. 简述数字签名的用途和基本流程。

34. 什么是数字签名？什么是消息认证？

35. 美国数字签名标准方案具体内容是什么？

36. 要将明文 $M$ 由$A_1$ 并附有$A_1$、$A_2$、…、$A_i$、…、$A_n$ 的依次签名发往B。设 $PK_{A_i}$ 和 $SK_{A_i}$ 分别为 $A_i$ 的公开密钥和私有密钥，在签名时要求每一位签名者只验证其前一位签名者的签名；如果验证通过，则在此基础上加上自己的签名，否则终止签名；最后一位签名者在签名完成后将最终信息和签名一起发送出去。每一位签名者都可以推算出前一位签名者和后一位签名者并且知道他们的公开密钥。试设计该多人签名算法。

37. 查阅相关资料，比较各种数字签名算法的优缺点。

38. 可信第三方有些什么作用？

39. 简述一个成功的SET交易的标准流程。

40. 电子支付中有哪些安全需求？

41. SET（安全电子交易）中有哪些关键技术？

# 第 3 章　身份认证与访问控制

信息系统中的一切活动都是由访问行为所引起的。为了系统的安全，需要对访问进行管制约束。访问涉及两个方面：主体（通常指用户）和客体（也称为资源，即数据）。身份认证（identity authentication）是指对于主体合法性的认证；访问控制（access control）是指对于主体的访问行为进行授权（authorization）的过程。

如果认为一个信息系统有一个入口，则身份认证就是在信息系统的入口进行的身份检查；而访问控制则规定访问者进入系统以后可以对哪些资源分别进行什么样的访问操作。两者之间的关系如图 3.1 所示。

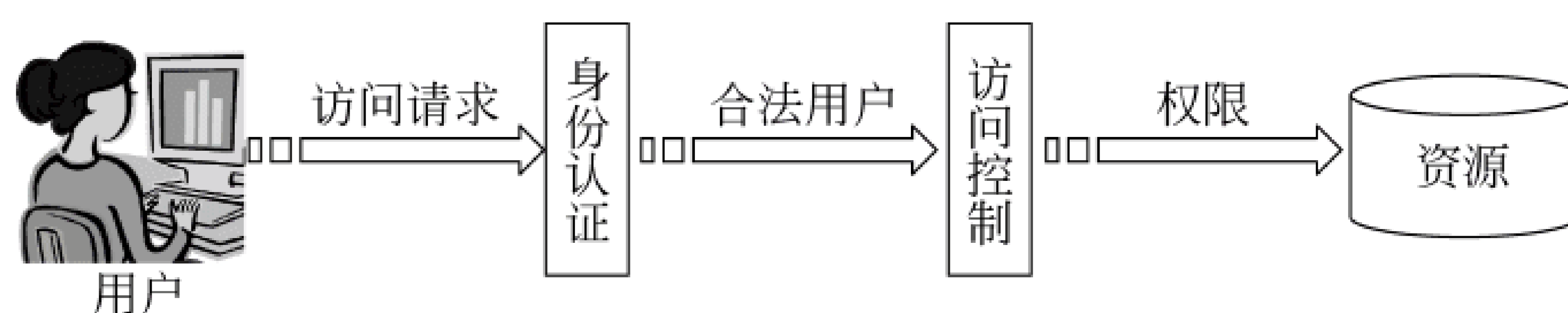



图 3.1　用户对资源访问的过程

## 3.1　基于凭证比对的身份认证

身份认证是信息系统安全的第一道屏障，用于检验主体身份的合法性，用它控制哪些用户能够登录到系统（服务器）并获取系统资源，控制准许用户进入的时间和准许他们在哪台计算机上访问。

最基本的身份认证是需要用户提供能代表身份的凭证与系统中存储的凭证进行比对。用于比对的身份凭证可分为下列 3 种。

（1）用户所知道的秘密，如口令（password）、个人识别号（PIN）和密钥等。

（2）用户所拥有的信物，如信用卡、IC 卡、USB Key、印章和证件等。

（3）用户自身的特征，如笔迹、步态、声音、指纹、虹膜纹和唇纹等。

身份认证可以是一方对另一方的认证，也可以是双方的互相认证。前者称为单向认证，后者称为双向认证。

### 3.1.1　生物特征身份认证

生物特征身份凭证一般采用用户固有的生物特征和行为特征，要求这些具有唯一性和永久性。下面介绍几种主要的生物身份凭证及其验证方法。

#### 1. 指纹

指纹是历史最为悠久的生物身份凭证。据著名指纹专家刘持平先生论证，早在 7000 年前我们的祖先就开始进行指纹识别的研究。到了春秋战国时代，手印检验不仅广泛应用于

政府和民间的书信与邮件往来之中，并已经开始用于侦讯破案之中。

指纹是一种十分精细的拓扑图形。如图 3.2 所示，一枚指纹不足方寸，上面密布着 100～120 个特征细节，这么多的特征参数组合的数量达到 640 亿种（英国学者高尔顿提出的数字）。并且由于它从胎儿 4 个月时生成后保持终生不变，因此，用它作为人的唯一标识，是非常可靠的。

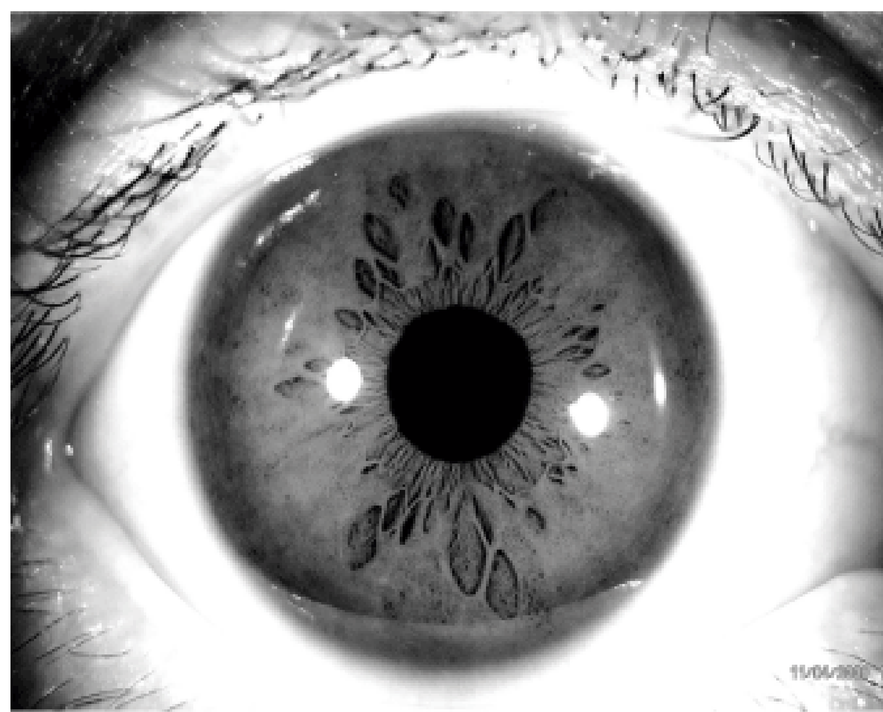


图 3.2　指纹的细节特征

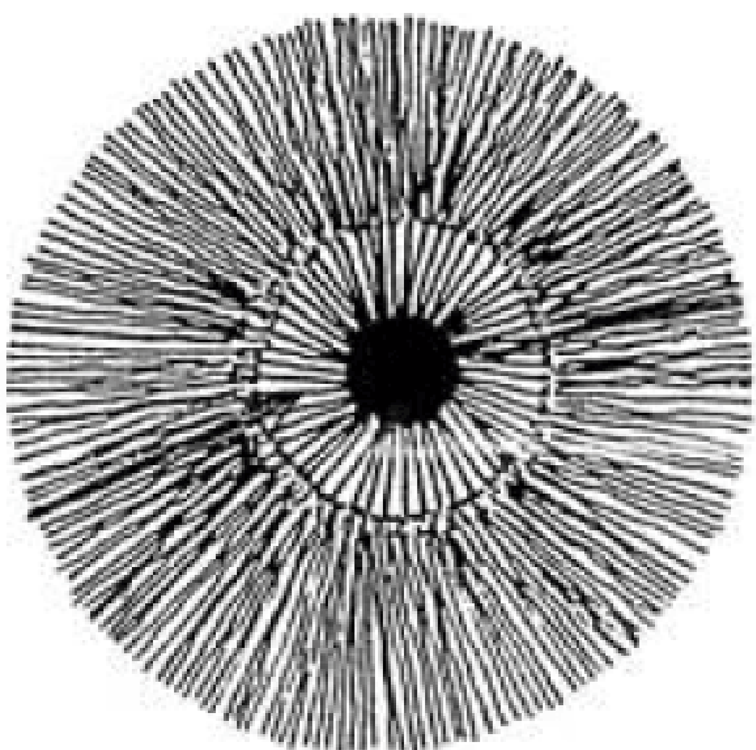
指纹识别主要涉及 4 个过程：读取指纹图像、提取指纹特征、保存数据和比对。目前已经开发出计算机指纹识别系统，可以比较精确地进行指纹的自动识别。

**2. 虹膜**

虹膜是位于眼睛黑色瞳孔与白色巩膜之间的环形部分（见图 3.3（a））。它在总体上呈由里向外的放射状结构（见图 3.3（b）），并包含许多相互交错的类似斑点、细丝、冠状、条纹、隐窝等形状的细微特征。这些细微特征信息也被称为虹膜的纹理信息，主要由胚胎发育环境的差异决定，因此对每个人都具有唯一性、稳定性和非侵犯性。



(a) 虹膜位置



(b) 虹膜结构

图 3.3　眼睛与虹膜

虹膜识别系统主要由虹膜图像采集装置、活体虹膜检测算法、特征提取和匹配几个模块组成。

**3. 面像**

采用面像作为身份凭证的识别系统包括两个技术环节：面像检测和面像识别。

1）面像检测

面像检测主要实现面像的检测和定位，即从输入图像中找到面像及面像的位置，并将人脸从背景中分割出来。现有的面像检测方法可以分为 3 类。

（1）基于规则的面像检测：总结了特定条件下可用于检测面像的知识（如脸型、肤色等），并把这些知识归纳成指导面像检测的规则。

（2）基于模板匹配的面像检测：首先构造具有代表性的面像模板，通过相关匹配或其他相似性度量检测面像。

（3）基于统计学习的面像检测：主要利用面部特征点结构灰度分布的共同性来检测面像。

2）面像识别

面像识别由两个过程组成。

（1）面像样本训练：提取面像特征，形成面像特征库。

（2）识别：用训练好的分类器将待识别面像的特征同特征库中的特征进行匹配，输出识别结果。

**4. 声纹**

声纹鉴定是以人耳听辨的声纹为基础，不仅关注发音人的语音频谱等因素，还充分挖掘说话人语音流中的各种特色性事件和表征性特点，如由方言背景确定的地域性，发音部位变化、内容以及发音速度和强度确定的发音人的年龄、性格和心态等。

在计算机处理时，常常将人类声纹特征分为 3 个层次。

（1）声道声学层次：在分析短时信号的基础上，抽取对通道、时间等因素的不敏感特征。

（2）韵律特征层次：抽取独立于声学、声道等因素的超音段特征，如方言、韵律和语速等。

（3）语言结构层次：通过对语音信号的识别，获取更加全面和结构化的语义信息。

声纹识别系统主要包括两部分。

（1）特征提取：选取唯一表现说话人身份的有效且可靠的特征。

（2）模式匹配：对训练和识别时的特征模式进行相似性匹配。

但是，目前还没有证实它的唯一性。

**5. 其他**

其他可用于身份识别的生物特征还有步态、笔迹、签名、颅骨外形、视网膜、唇纹、DNA、按键特征、耳朵轮廓、体温图谱、掌形和足迹等。

### 3.1.2 静态口令

口令通常是作为用户账号补充部分向系统提交的身份凭证。一般说来，用户账号是公开的。当用户向系统提交了账号以后，还需要提交保密形式的凭证——口令，供系统鉴别用户的真实性，以防止非法使用用户账号登录。所以，用户只有向系统输入口令，通过了系统的验证后，才能获得相应的权限。

口令是使用度最高的一类身份认证，它简单、易用，效率很高，但它也是极为脆弱、容易攻击的认证方式。

**1. 口令失密及其对策**

口令是较弱的安全机制。从责任的角度看，用户和系统管理员都对口令的失密负有责任，或者说用户和系统管理员两方都有可能造成口令失密。从失密的途径看，有众多的环

节可以造成口令失密，或者说，攻击者可以从下面一些途径进行口令攻击。

（1）猜测和发现口令。

① 常用数据猜测，如家庭成员或朋友的名字、生日、球队名称、城市名、身份证号码、电话号码和邮政编码等。

② 字典攻击：按照字典序进行穷举攻击。

③ 其他，如望远镜窥视等。

（2）电子监控。

在网络或电子系统中，被电子嗅探器和监控器窃取。

（3）访问口令文件。

① 在口令文件没有强有力保护的情形下，下载口令文件。

② 在口令文件有保护的情况下，进行蛮力攻击。

（4）通过社交。

如通过亲情、收买或引诱，获取别人的口令。

（5）垃圾搜索。

收集被攻击者的遗弃物，从中搜索被疏忽丢掉的写有口令的纸片或保存有口令的盘片。

（6）用蠕虫记录用户输入的口令。

蠕虫可以根据用户按键的位置记录用户输入的口令。

#### 2. 口令的安全保护

口令一旦失密或被破解，该用户的账号就不再受到保护，攻击者就可以大摇大摆地进入系统。因此，口令的保护是用户和系统管理员都必须重视的工作。下面从几个方面考虑口令的安全。

1）正确选取选取口令字

选取口令应遵循以下原则。

（1）扩大口令的字符空间。口令字符空间越大，穷举攻击的难度就越大。一般地，不要仅限于使用 26 个大写字母，可以扩大到小写字母、数字等计算机可以接受的字符空间。

（2）选择长口令。口令越长，破解需要的时间就越长，一般应使口令位数大于 6 位。

（3）使用随机产生的口令，避免使用弱口令（有规律的口令）和容易被猜测的口令，如家庭成员或朋友的名字、生日、球队名称和城市名等。

（4）使用多个口令，在不同的地方不要使用相同的口令。

2）正确地使用口令

（1）缩短口令的有效期。口令要经常更换，最好使用动态的一次性口令。

（2）限制口令的使用次数。

（3）限制登录时间，如属于工作关系的登录，把登录时间限制在上班时间内。

3）安全地保存指令

口令的存储不仅是为了备忘，更重要的是系统要在检测用户口令时进行比对。直接明文存储口令（写在纸上或直接明文存储在文件或数据库中）最容易泄密。较好的方法是将每一个用户的系统存储账号和哈希值存储在一个口令文件中。当用户登录时，输入口令后，系统计算口令的哈希码，并与口令文件中的哈希值比对，若相等，则允许登录，否则拒绝登录。

4）口令审计

系统管理员除对用户账户要按照权限等加以控制外，还要对口令的使用在以下几个方面进行审计。

（1）最小口令长度。

（2）强制修改口令的时间间隔。

（3）口令的唯一性。

（4）口令过期失效后允许入网的宽限次数。如果在规定的次数内输入不了正确口令，则认为是非法用户的入侵，应给出报警信息。

5）增加口令认证的信息量

例如，在认证过程中，随机地提问一些与该用户有关，并且只有该用户才能回答的问题。

6）使用软键盘输入口令

软键盘是一种显示在屏幕上的键盘，可供用户用鼠标选择单击进行输入。如图 3.4 所示，软键盘上的键的布局可以是随机的，这样就能有效地防止木马通过对按键位置的记录，窃取用户密码。




图 3.4　某银行的客户端登录软键盘

**3. 批量登录攻击与验证码**

验证码也称为全自动区分计算机和人类的图灵测试（Completely Automated Public Turing test to tell Computers and Humans Apart，CAPTCHA），是一种区分用户是计算机还是人的公共全自动技术，其目的是有效防止某一个特定注册用户用特定程序暴力破解方式进行不断的登录尝试。其具体方法是强迫用户在登录时必须要人工进行一些工作。基本方法是要用户从模糊的图形之中辨认出隐藏在其中的一些信息。用户只有将正确的答案与账户和口令一起发送，才能注册。这就使得攻击者使用程序自动注册成为不可能。

在形式上，验证码可以是数字、字母、文字、图片、广告以及问题等。

验证码主要用来控制注册或登录的时间和节奏不能太快。用户登录时，验证码根据时间周期随机生成，用户在一定的时间周期内必须从图片中人工找出所隐藏的信息，输入验证码，提交服务器系统验证，验证成功才能登录。两次登录之间有一个验证生存期（一般为 30s）。

强制人为干预的另一种方法是通过手机传送验证码。

### 3.1.3 动态口令

#### 1. 动态口令的概念

动态口令也称为一次性口令（One-Time Password，OTP），是最安全的口令。它是根据专门的算法生成一个不可预测的随机数字组合，每个密码只能使用一次，目前被广泛运用在网银、网游、电信运营商、电子商务、企业等应用领域。

动态口令有如下优势。

（1）可以提供给最终用户安全访问企业核心信息的手段。

（2）可以降低与密码相关的 IT 管理费用。

（3）是一种无须记忆的复杂密码，降低了遗忘密码的概率。

#### 2. 动态令牌的类型

为了安全，动态口令不是在网络上直接生成，也不是由系统直接从网络上发给用户，而是通过其他渠道或生成器提供给用户。这些用于生成动态口令终端通常称为“令牌”。 目前主流令牌有短信密码、手机令牌、硬件令牌和软件令牌 4 种。

1）短信密码

短信密码以手机短信形式请求包含 6 位或更多随机数的动态口令，身份认证系统以短信形式发送随机的 6/8 位密码到客户的手机上，客户在登录或者交易认证的时候输入此动态口令，从而确保系统身份认证的安全性。

2）手机令牌

手机令牌是一种手机客户端软件，它每隔 30s 产生一个随机 6 位动态密码，口令生成过程不产生通信及费用，具有使用简单、安全性高、低成本、无须携带额外设备、容易获取、无物流等优势，手机令牌是 3G 时代动态密码身份认证发展的趋势。手机令牌有 J2ME、iPhone、Android 和 Windows Mobile 6 版本，可以广泛应用在网络游戏、互联网等用户基数大的领域，手机令牌的使用将大大减小动态密码服务管理及运营成本，方便用户。

3）硬件令牌

硬件令牌往往是一个钥匙扣大小的轻巧器具，上有显示屏，可以显示随机密码。它每 60s 变换一次动态口令，动态口令一次有效，它产生 6/8 位动态数字。图 3.5 是一款硬件令牌。




图 3.5 一款硬件令牌

4）软件令牌

软件令牌是通过软件生成随机密码。

#### 3. 动态口令技术分类

动态口令技术主要分两种：同步口令技术和异步口令技术（挑战-应答方式）。其中的同步口令技术又分为时间同步口令和事件同步口令两种。

1）同步口令

（1）时间同步口令。

时间同步口令基于令牌和服务器的时间同步，并且采用国际标准时间，一般每 60s 产生一个新口令。为了保持服务器与令牌的同步，一方面，要求服务器要能够十分精确地保持正确的时钟，对令牌的晶振频率也有严格的要求；另一方面，由于令牌的工作环境不同，在磁场、高温、高压、震荡、浸水等情况下易发生时钟脉冲的不确定偏移和损坏，因此在每次进行认证时，服务器端将会检测令牌的时钟偏移量，不断微调自己的时间记录。

（2）事件同步口令。

基于事件同步的令牌是通过某一特定的事件次序及相同的种子值作为输入，通过哈希算法运算出一致的密码。其整个工作流程与时钟无关，不受时钟的影响，令牌中不存在时间脉冲晶振。但由于其算法的一致性，其口令是预先可知的，通过令牌，可以预先知道今后的多个密码，故当令牌遗失且没有使用 PIN 码对令牌进行保护时，存在非法登录的风险。因此对于 PIN 码的保护是十分必要的。

2）异步口令

异步口令不需要令牌和服务器之间同步，因而降低了对应用的影响，极大地提高了系统的可靠性。它的主要技术是采用了挑战/应答（challenge-response）方式。

基于挑战/应答方式的身份认证系统在每次认证时，认证服务器端都给客户端发送一个不同的“挑战”字串，客户端程序收到这个“挑战”字串后，做出相应的“应答”，具体过程如下。

（1）客户向认证服务器发出请求，要求进行身份认证。

（2）认证服务器从用户数据库中查询用户是否是合法的用户，若不是，则不做进一步处理。

（3）认证服务器内部产生一个随机数，作为“提问”，发送给客户。

（4）客户将用户名字和随机数合并，使用单向哈希函数（例如 MD5 算法）生成一个 6/8 位的随机数字字节串作为应答，口令一次有效。

（5）认证服务器将应答串与自己的计算结果比较，若两者相同，则通过一次认证；否则，认证失败。

（6）认证服务器通知客户认证成功或失败。

以后的认证由客户不定时地发起，过程中没有了客户认证请求这一步。这个过程增加了用户操作的复杂度，因此两次认证的时间间隔不能太短，否则就给网络、客户和认证服务器带来太大的开销；也不能太长，否则不能保证用户不被他人盗用 IP 地址，一般定为 1～2min。

#### 4. 智能卡

智能卡（smart card）是一种集成电路卡，它内置有处理器，可以存储用户的个性化的秘密信息，并提供硬件保护措施和加密算法。进行认证时，用户输入自己的 PIN，先由智能卡进行认证。认证成功后，即可读出卡中的秘密信息，进而进行与主机间的认证。

**5. 基于挑战/应答的双因子 USB Key**

一种常用的智能卡是基于挑战/应答的双因子 USB Key（见图 3.6）。双因子认证是指用户必须具备两个必要因素（如 PIN 和 USB Key），才可以登录系统。这样，即使 PIN 暴露，只要 USB Key 不同时被获得 PIN 的人掌握，用户的合法身份就不会被假冒；或者 USB Key 遗失，但没有掌握用户的 PIN，用户的合法身份也不会被假冒。



图 3.6　一款 USB Key

## 3.2　基于密钥分发的身份认证

密钥是加密的工具，但由于密钥的私密性，使它也具有凭证的某些特性。在网络环境下，密钥分发是通过握手过程进行的，这个过程要遵守某种规则，这些称为认证协议或算法。在这个过程中也有了身份认证的作用。需要注意的是，“身份”不仅包含真实性，还包含时效性，即此刻的 A 不是彼刻的 A，只有确认了这一点，才能有效地防止抵赖行为。

### 3.2.1　公钥加密认证协议

公钥加密认证协议是基于公钥加密体制分配会话密钥过程实现的。下面介绍几种认证协议。

**1. 相互认证协议**

（1）一个通过认证服务器 AS 的身份认证协议如下。

① A→AS：$\mathrm{ID_A} \| \mathrm{ID_B}$。

② AS→A：$E_{\mathrm{SK_{AS}}}[\mathrm{ID_A} \| \mathrm{PK_A} \| T] \| E_{\mathrm{SK_{AS}}}[\mathrm{ID_B} \| \mathrm{PK_B} \| T]$。

③ A→B：$E_{\mathrm{SK_{AS}}}[\mathrm{ID_A} \| \mathrm{PK_A} \| T] \| E_{\mathrm{SK_{AS}}}[\mathrm{ID_B} \| \mathrm{PK_B} \| T] \| E_{\mathrm{PK_B}}[E_{\mathrm{SK_A}}[K_\mathrm{S} \| T]]$。

这个协议需要各方时钟同步。

（2）一个通过 KDC 的身份认证协议如下。

① A→KDC：$\mathrm{ID_A} \| \mathrm{ID_B}$。

② KDC→A：$E_{\mathrm{SK_{AU}}}[\mathrm{ID_B} \| \mathrm{PK_B}]$（$\mathrm{SK_{AU}}$ 是 KDC 的私钥）。

③ A→B：$E_{\mathrm{PK_B}}[N_\mathrm{A} \| \mathrm{ID_A}]$（$N_\mathrm{A}$ 是 A 选择的一次性随机数）。

④ B→KDC：$\mathrm{ID_B} \| \mathrm{ID_A} \| E_{\mathrm{PK_{AU}}}[N_\mathrm{A}]$（$\mathrm{PK_{AU}}$ 是 KDC 的公钥）。

⑤ KDC→B：$E_{\mathrm{SK_{AU}}}[\mathrm{ID_A} \| \mathrm{PK_A}] \| E_{\mathrm{PK_B}}[E_{\mathrm{SK_{AU}}}[N_\mathrm{A} \| K_\mathrm{S} \| \mathrm{ID_B}]]$（$K_\mathrm{S}$ 是 KDC 为 A、B 分配的一次性会话密钥）。

⑥ B→A：$E_{\mathrm{PK_A}}[E_{\mathrm{SK_{AU}}}[N_\mathrm{A} \| K_\mathrm{S} \| \mathrm{ID_B}] \| N_\mathrm{B}]$。

⑦ A→B：$E_{K_\mathrm{S}}[N_\mathrm{B}]$。

这个协议中使用了一次性随机数，所以不再要求各方时钟的同步。但是，这个协议不

能抵御攻击者对 A 的假冒。请读者设法为此改进这个协议。

**2. 单向认证协议**

简单地说，认证协议主要有两种作用：提供机密性和认证性。在公钥加密体制中，这些功能要看是否为对方提供公钥。

（1）发送方知道接收方的公钥，才有可能实现机密性保护。例如，下面的协议仅提供机密性：

A→B：$E_{PK_B}[K_S] \| E_{K_S}[M]$（$K_S$为 A 向 B 发送的一次通信密钥）

（2）接收方知道发送方的公钥，才有可能实现认证性保护。例如，下面的协议仅提供认证性：

A→B：$M \| E_{SK_A}[H(M)]$

这时，为了使 B 确信 A 的公钥的真实性，A 还要向 B 发送公钥证书：

A→B：$M \| E_{SK_A}[H(M)] \| E_{SK_{AS}}[T \| ID_A \| PK_A]$（$SK_{AS}$为认证服务器的公钥，$E_{SK_{AS}}[T \| ID_A \| PK_A]$是 AS 给 A 签署的公钥证书）。

（3）发送方和接收方互相知道对方的公钥，则既可提供机密性又可提供认证性，例如：

A→B：$E_{PK_B}[M \| E_{SK_A}[H(M)]]$

这时，为了使 B 确信 A 的公钥的真实性，A 还要向 B 发送公钥证书：

A→B：$E_{PK_B}[M \| E_{SK_A}[H(M)]] \| E_{SK_{AS}}[T_S \| ID_A \| PK_A]$

### 3.2.2 单钥加密认证协议

**1. 相互认证协议 Needham-Schroeder**

1）Needham-Schroeder 协议过程

如 2.1.5 节所述，通过 KDC 进行单密钥分配，通常采用图 2.6 所示的方法。这个过程由 5 步组成。

① A→KDC: $ID_A \| ID_B \| N_A$（A 和 KDC 请求与 B 加密通信）。

② KDC→A：$E_{K_A}[K_S \| ID_B \| N_A \| E_{K_B}[K_S \| ID_A]]$（A 获得 $K_S$）。

③ A→B：$E_{K_B}[K_S \| ID_A]$（B 安全地获得 $K_S$）。

④ B→A：$E_{K_S}[N_B]$（B 知道 A 已掌握 $K_S$，用加密 $N_B$ 向 A 示意自己也获得 $K_S$）。

⑤ A→B：$E_{K_S}[f(N_B)]$。

这个协议被称为 Needham-Schroeder 协议。在这个协议中，前 3 步是 KDC 分发密钥，第④、⑤两步是一次握手过程，即 B 认证 A 的过程：当在第⑤步中 B 能正确收到自己在第④步发出的 $N_B$ 时，就可以证明 A 是当前的通话对象，因为自己在第③步获得的 $K_S$ 是“新鲜”的，而非攻击者截获的前一次执行通话时用过的 $K_S$ 的重放。但是，若攻击者已经获得旧会话密钥 $K_S$，并冒充 A 向 B 重放第③步的消息，就可以欺骗 B 使用旧 $K_S$ 会话，接着截获第④步 B 的询问，再冒充 A 对 B 应答。这样就能冒充 A 向 B 发送假消息，使抗抵赖保

护失败。

2）Needham-Schroeder 协议改进

改进的办法是在第②步和第③步中加上一个时间戳，即

② KDC→A：$E_{K_A}[K_S \| ID_B \| T \| \ E_{K_B}[K_S \| ID_A \| T]]$。

③ A→B：$E_{K_B}[K_S \| ID_A \| T]$。

这样，A 和 B 都可以利用当前时间对 $T$ 进行检查，以确定 $K_S$ 是否陈旧的。但是，使用这个协议的前提是 A 和 B 的时钟完全同步。若由于系统故障或存在计时误差，就会被攻击者利用时间差进行重放攻击。

重放攻击（replay attacks）又称为重播攻击、回放攻击或新鲜性攻击（freshness attacks），是指攻击者发送一个目的主机已接收过的包，用欺骗手法破坏认证的正确性。

3）Needham-Schroeder 协议再改进

克服上述缺陷的方法是将 Needham-Schroeder 协议进一步改进为以下过程。

① A→B：$ID_A \| N_A$。

② B→KDC：$ID_B \| N_B \| \ E_{K_B}[ID_A \| N_A \| T_B]$。

③ KDC→A：$E_{K_A}[ID_B \| N_A \| K_S \| ID_A \| T_B] \| E_{K_B}[ID_A \| K_S \| T_B] \| N_B$。

④ A→B：$E_{K_B}[ID_A \| K_S \| T_B] \| E_{K_S}[N_B]$。

这个协议的执行过程如图 3.7 所示。

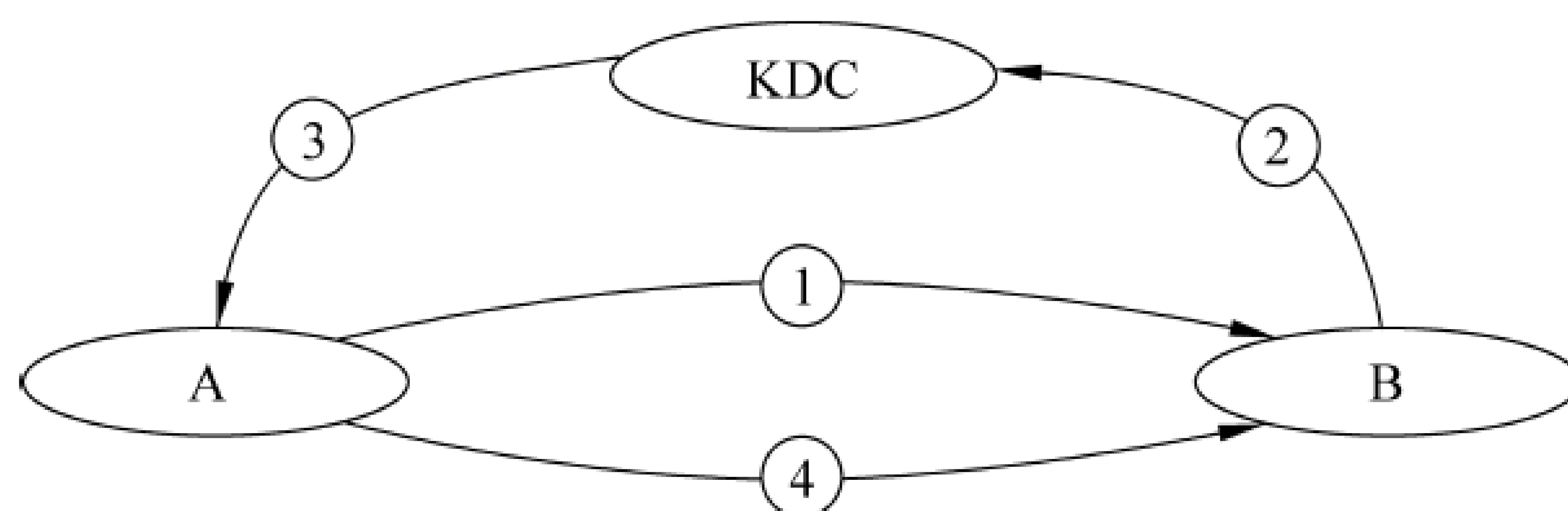



图 3.7　进一步改进的 Needham-Schroeder 协议

分析这个进一步改进的 Needham-Schroeder 协议可以看出以下几点。

在第①步中，A 将 $ID_A$ 和 $N_A$ 以明文传送给 B；在第②步中，B 用自己和 KDC 共享的主密钥对 $ID_A$ 和 $N_A$ 加密传送给 KDC；在第③步中，KDC 对从 B 传来的信息解密，再用 KDC 与 A 共享的主密钥，将 $K_S$ 和 $N_A$ 一同加密传回 A。A 验证了 $N_A$ 就可以知道，B 已经收到了 A 在第①步中发送的消息，同时也确信 $K_S$ 是新鲜的。

在第②步中，B 将 $ID_B$ 和 $N_B$ 以明文传送给 KDC，经第③步由 KDC 将 $N_B$ 传送给 A，再由 A 用 $K_S$ 加密 $N_B$ 传回 B。同 $N_A$ 的作用一样，$N_B$ 用来保证 B 收到的 $K_S$ 是新鲜的。

在第②步中，B 发出的 $T_B$ 是 B 建议的证书截止时间，它是 B 根据自己的时钟确定的，不要求各方之间同步。

$E_{K_B}[ID_A \| N_A \| T_B]$经 KDC 传送给 A，由 A 留作以后认证的证据，并可以在有效时间范围内，不借助认证服务器（KDC）而是通过以下几步实现双方的新认证。

① A→B：$E_{K_B}[ID_A \| K_S \| T_B]$，$N'_A$。

② B→A：$N'_{B}$，$E_{K_S}$ [$N'_{A}$]。

③ A→B：$E_{K_S}$[$N'_{B}$]。

这里，B 通过 $T_B$ 检验证据是否过时，而新产生的随机数 $N'_A$ 和 $N'_B$ 可以用来保证没有重放攻击。

**2. 单向认证协议**

对于单向保密通信的特点，在 Needham-Schroeder 协议中去掉第④步和第⑤步，就成为能满足单向通信两个基本要求的单向认证协议。

① A→KDC: $ID_A \| ID_B \| N_A$。

② KDC→A：$E_{K_A}[K_S \| ID_B \| N_A \| E_{K_B}[K_S \| ID_A]]$。

③ A→B：$E_{K_B}[K_S \| ID_A] \| E_{K_S}[M]$。

这个协议提供了对于发送方 A 的认证，保证只有 B 才能阅读报文。但是，它不能防止重放攻击。为此，可以使用时间戳。不过由于电子邮件处理的延迟性，时间戳的作用有限。

### 3.2.3 Kerberos 认证系统

Kerberos 是 MIT（麻省理工学院）的一种基于 Needham-Schroeder 算法的网络认证系统，其设计目标是通过对称密钥系统为客户/服务器应用程序提供强大的认证服务。Kerberos 的名称来自希腊神话中一种有 3 个脑袋的地狱守门狗。设计者采用这个名字是想给网络大门提供如下 3 种守护。

（1）认证（authentication）。

（2）清算（accounting）。

（3）审计（audit）。

**1. Kerberos 的计算环境**

Athena 认为，Kerberos 计算环境由大量的匿名工作站和相对较少的独立服务器组成。服务器提供文件存储、打印、邮件等服务，工作站主要用于交互和计算。在此环境中存在如下 3 种威胁。

（1）用户可以访问特定的工作站并伪装成该工作站用户。

（2）用户可以改动工作站的网络地址，伪装成其他工作站。

（3）用户可以根据交换窃取消息，并使用重放攻击来进入服务器。

在整个网络中，除了 Kerberos 服务器外，其他都是危险区域，任何人都可以在网络上读取、修改和插入数据。作为一种可信任的第三方认证服务，在这样的环境下，Kerberos 认证过程的实现不依赖于主机操作系统的认证，无须基于主机地址的信任，不要求网络上所有主机的物理安全，并假定网络上传送的数据包可以被任意地读取、修改和插入数据。

**2. Kerberos 系统组成**

一个完整的 Kerberos 主要由如下几个部分组成。

1）两个服务对象

（1）客户（client）：发起认证服务方。
（2）服务器（server）：接受客户端的请求，对数据库进行操作。

2）两类凭证

在 Kerberos 中使用两类凭证。

（1）票证（Ticket Granting Ticket，TGT），也称为入场券：用来安全地在认证服务器和用户请求的服务之间传递信息，内容包括：①用户的身份；②下一阶段通信双方使用的临时加密密钥——会话密钥（session key）；③时间标记（timestamp），用于检测重放攻击（replay attack）。入场券一旦生成，在其生存期内可以被用户多次使用来申请同一个服务器的服务。

（2）鉴别码（authenticator）：用来作为认证凭证的一段加密文字，用来提供信息与入场券中的信息进行比较，一起保证发出入场券的用户就是入场券中指定的用户，以防止攻击者再次使用同一凭证。其内容包括校验和、子密钥、序列号和身份认证数据。它们只能在一次服务请求中使用，每当用户向服务器申请服务时，必须重新生成 Authenticator。

两种凭证均使用私有密钥加密，但加密的密钥不同。

3）两个库

（1）Kerberos 数据库（中心数据库）：记载了每个 Kerberos 用户的名字、用户口令、私有密钥和截止信息（记录的有效时间，通常为几年）等重要信息。

（2）Kerberos 应用程序库：应用程序接口，包括创建和读取认证请求，以及创建 safe message 和 private message 的子程序。

4）两个服务器

为了减轻每个服务器的负担，Kerberos 把身份认证的任务集中在身份认证服务器上。Kerberos 的认证服务任务被分配给两个相对独立的服务器。

（1）认证服务器（Authenticator Server，AS）：存放一个 Kerberos 数据库的只读副本，生成会话密钥，验证用户身份。当一个用户登录到一个企业内部网请求访问内部服务器时，AS 将根据中心数据库存储的用户密码生成一个 DES 加密密钥，对一个入场券（TicketTGS）进行加密。这个入场券是提供给 TGS 的。

（2）票据分配服务器（Ticket Granting Server，TGS）：也称为入场券许可服务器，用于发放身份证明票据（凭证）。当用户要访问某个服务器 S 时，TGS 就会查找中心数据库中的存取控制表，以确认该用户是否已经授权使用该服务器。确认后，会生成一个新的凭证（TicketS，相当于手牌）。这个新的凭证包含有与服务器相关的密钥和加密后的入场券（TicketTGS）。

**3．Kerberos 系统认证使用的信息**

（1）在认证过程中使用以下身份识别码。

$ID_C$：客户（工作站）标识。

$ID_T$：TGS 标识。

$ID_S$：服务器标识。

（2）Kerberos 是采用对称密钥加密算法来通过第三方 KDC 实现身份认证。在认证过程中使用如下密钥。

$K_C$：C 的用户密钥，由 C 上的用户口令导出，可与 AS 共享。

$K_S$：S 的用户密钥，可与 TGS 共享。

$K_T$：TGS 的用户密钥，AS 与 TGS 共享。

$K_{CT}$：会话密钥，C 与 TGS 共享。

$K_{CS}$：会话密钥，C 与 S 共享。

用户密钥属于长效密钥（long-term key）；会话密钥属于短效密钥（short-term key），仅用于一次会话。

（3）在认证过程中使用如下数据。

$AD_C$：C 的网络地址。

$TS_i$：第 $i$ 个时间戳。

$Lifetime_i$：第 $i$ 个有效期间。

（4）在认证过程中获取如下凭证。

$Ticket_S$：服务器入场券。

$Ticket_{TGS}$：TGS 的入场券。

$Authenticator_C$：用户生成的最终认证信息。

### 4．Kerberos 同域认证过程

Kerberos 的同域认证过程如图 3.8 所示，分为 3 个阶段 6 个步骤。

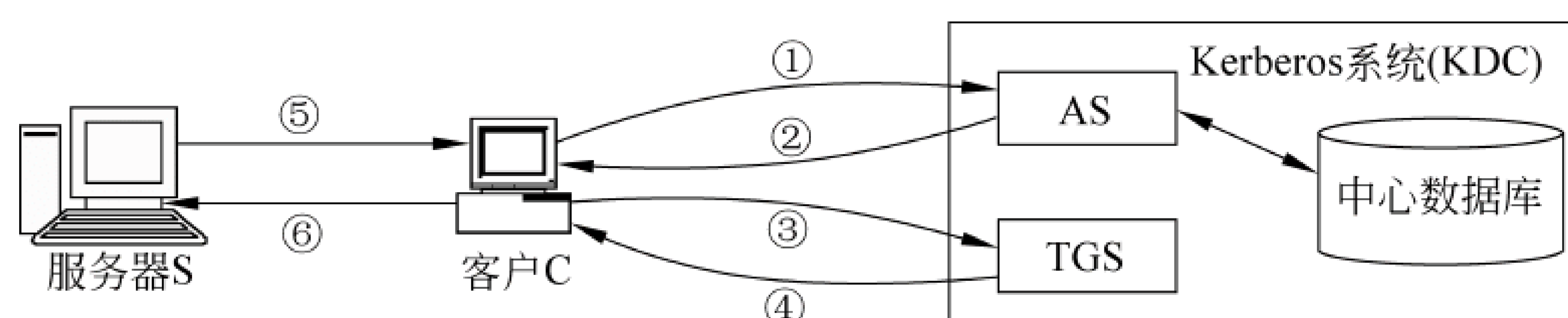



图 3.8　Kerberos 同域认证

（1）认证服务交换，客户从 AS 取得入场券。

①：客户向 AS 发出访问 TGS 请求（用 $TS_1$ 防止回放攻击）：

$C \rightarrow AS$：$E_{K_C}[ID_C \| ID_T \| TS_1]$。

②：AS 向 C 发放入场券——TGS 许可票证 $Ticket_{TGS}$：

$AS \rightarrow C$：$E_{K_C}[K_{CT} \| ID_T \| TS_2 \| Lifetime_2 \| Ticket_{TGS}]$。

其中：

$K_{CT}$ 为供 C 与 TGS 共享的会话密钥。

$Ticket_{TGS} = E_{K_T}[K_{CT} \| ID_C \| AD_C \| ID_T \| TS_2 \| Lifetime_2]$。

（2）入场券许可服务交换，C 用入场券换取 TGS 服务许可凭证。

③：C 向 TGS 发出请求，内容包括服务器识别码、入场券和一个认证符：

C→TGS： $E_{K_{CT}}$ [$ID_S$ || $Ticket_{TGS}$ || $Authenticator_C$]。

其中：

$Ticket_{TGS}$= $E_{K_T}$ [$K_{CT}$ || $ID_C$ || $AD_C$ || $ID_T$ || $TS_2$ || $Lifetime_2$]（提交入场券供 TGS 认证）。

$Authenticator_C$= $E_{K_{CT}}$ [$ID_C$ || $AD_C$ || $TS_3$]（向 TGS 提供自己的鉴别码）。

④：TGS 验证，向 C 发出服务许可凭证：

TGS→C： $E_{K_{CT}}$ [$K_{CS}$ || $ID_S$ || $TS_4$ || $Ticket_S$]。

其中：

$K_{CS}$ 为 C 与 S 的会话密钥。

$Ticket_S$= $E_{K_{TS}}$ [$K_{CS}$ || $ID_C$ || $AD_C$ || $ID_S$ || $TS_4$ || $Lifetime_4$]（C 无法解密，只能转交 S 认证）。

（3）客户和服务器相互认证，用户从服务器获取服务。

⑤：C 向服务器证明自己的身份（用 $Ticket_S$ 和 $Authenticator_S$）：

C→S： $E_{K_{CS}}$ [$Ticket_S$ || $Authenticator_C$]。

其中：

$Ticket_S$= $E_{K_{TS}}$ [$K_{CS}$ || $ID_C$ || $AD_C$ || $ID_S$ || $TS_4$ || $Lifetime_4$]。

$Authenticator_S$= $E_{K_{CS}}$ [$ID_C$ || $AD_C$ || $TS_5$]。

S 用 $Ticket_S$ 与 $Authenticator_S$ 对比，进行认证。

⑥：服务器向客户证明自己的身份：

S→C： $E_{K_{CS}}$ [$TS_5$+1]。

这个过程结束，客户 C 与服务器 S 之间就建立起了共享会话密钥，以便以后进行加密通信或交换新密钥。

### 5．Kerberos 异域认证

由于管理控制、政治经济和其他的因素，不太可能在世界范围内实现统一的 Kerberos 的认证中心，而每一个 Kerberos 的认证中心都具有或大或小的一定监管区域（Kerberos 的认证域），客户向本 Kerberos 的认证域以外的服务器申请服务的过程如图 3.9 所示，分为 7 步。

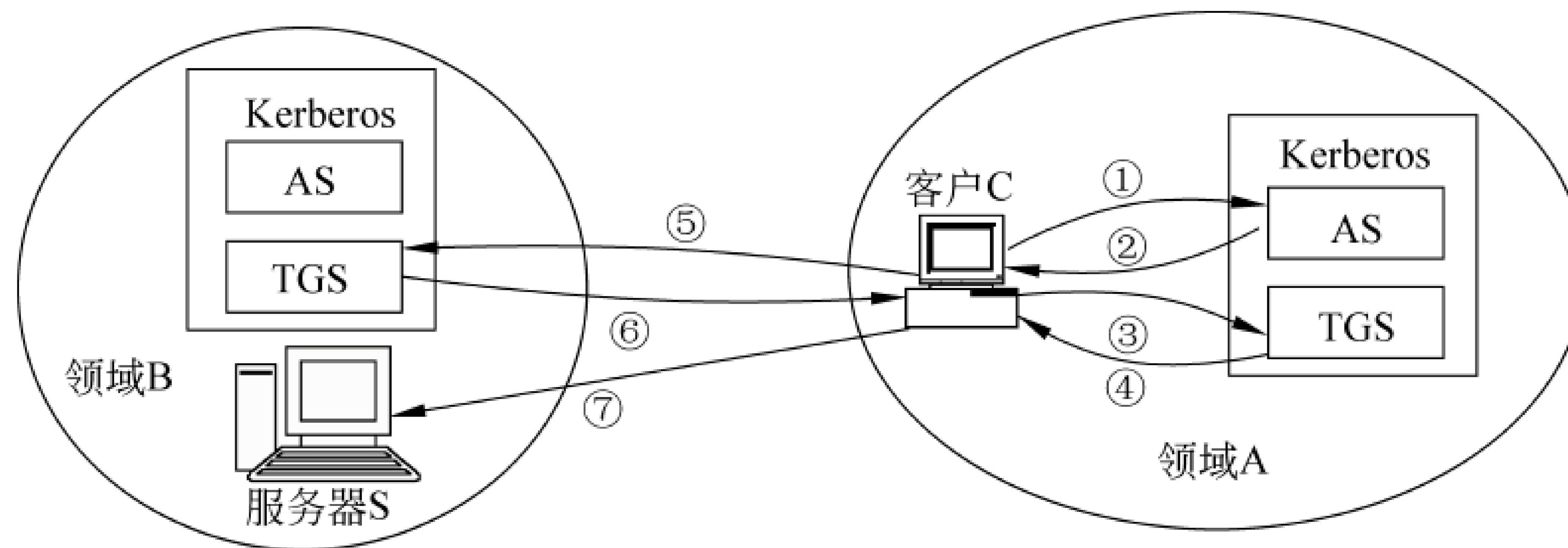



图 3.9　Kerberos 异域认证

① C→AS： $E_{K_C}$ [$ID_C$ || $ID_T$ || $TS_1$]。

② AS→C： $E_{K_C}$ [$K_{CT}$ || $ID_T$ || $TS_2$ || $Lifetime_2$ || $Ticket_{TGS}$]。

③ C→TGS： $E_{K_{CT}}$ [$ID_{TB}$ || $Ticket_{TGS}$ || $Authenticator_C$]（$ID_{TB}$ 为领域 B 的 $TRS_B$ 标识）。

④ TGS→C：$E_{K_{CT}}[K_{CTB} \| ID_{TB} \| TS_4 \| Ticket_{TB}]$（$K_{CTB}$ 为 C 与 $TRS_B$ 会话的密钥）。

⑤ C→$TGS_B$：$E_{K_{CTB}}[ID_{TB} \| Ticket_{TB} \| Authenticator_C]$。

⑥ $TGS_B$→C：$E_{K_{CTB}}[K_{CS} \| ID_{TB} \| \| TS_6 \| Ticket_{TB}]$。

⑦ C→S：$E_{K_{CS}}[Ticket_{TB} \| Authenticator_C]$。

## 3.3 基于数字证书的身份认证

### 3.3.1 数字证书

#### 1. 数字证书的特点

数字证书也称为数字身份证或数字 ID，其实就是一个特殊的计算机文件，这个计算机文件由权威机构——认证中心（Certificate Authority，CA）制作，是给网上用户的一组数字信息，包含用户身份信息、用户公开密钥、签名算法标识、证书有效期、证书序列号、颁证单位和扩展项等。数字证书有以下特点。

（1）它包含了身份信息，因此可以用于证明用户身份。

（2）它包含了非对称密钥，不但可用于数据加密，还可用于数据签名，保证通信过程的安全和不可抵赖。

（3）由于它是权威机构颁布的，因此具有很高的公信度。

有了数字证书之后，在网上通信的双方进行联系的第一步便是利用预装在浏览器中的安全认证软件和认证中心的公钥对通信对象的数字证书进行验证；验证无误后，才可使用认证中心传递的加密公钥进行加密通信。

#### 2. 基于数字证书的 USB Key

基于数字证书的智能卡，是用智能卡作为数字证书的存储介质，可以保证数字证书不被复制，并可以实现数字证书的所有功能。中国农业银行的 U 盾（见图 3.10）实际上是一种基于数字证书的 USB Key。它不仅可以履行数字证书的功能，进行双因子认证，还可以自动生成一个随机布局的软键盘供用户输入自己的 PIN，从多个角度保障客户账户安全。




图 3.10 中国农业银行的 U 盾

#### 3. 数字证书分类

常见的数字证书有以下几种。

（1）Web 服务器证书。用于 Web 服务器与用户浏览器之间建立安全连接通道。

（2）服务器身份证书。提供服务器身份信息、公钥和 CA 签名，用于确保与其他服务器或用户通信的安全。

（3）计算机证书。提供计算机的身份信息，确保与其他计算机通信的安全性。

（4）个人证书。提供证书持有人的个人身份信息、公钥和 CA 签名，用于在网络中标识个人身份。浏览器证书也是一种个人证书。

（5）安全电子邮件证书。提供证书持有者的电子邮件地址、公钥和 CA 签名，用于电子邮件的安全传递和认证。

（6）企业证书。提供企业的身份信息、公钥和 CA 签名，用于在网络中标识证书持有者的身份。

（7）代码签名证书。附加在软件代码中，用于证实软件的真实性，保护软件代码的完整性的数字证书。

#### 4. 认证中心

认证中心（CA）是可以信赖的第三方机构，具有如下一些功能。

（1）颁发证书。如密钥对的生成，私钥的保护等，并保证证书持有者应有不同的密钥对。

（2）管理证书。记录所有颁发过的证书以及所有被吊销的证书。

（3）用户管理。对于每一个新提交的申请，都要和列表中现存的标识名相比照，如出现重复，就予以拒绝。

（4）吊销证书。在证书有效期内使其无效，并发表 CRL。

（5）验证申请者身份。对每一个申请者进行必要的身份认证。

（6）保护证书服务器。证书服务器必须是安全的，CA 应采取相应措施保证其安全性。例如，加强对系统管理员的管理以及防火墙保护等。

（7）保护 CA 私钥和用户私钥。CA 签发证书所用的私钥要受到严格的保护，不能被毁坏，也不能非法使用。同时，根据用户密钥对的产生方式，CA 在某些情况下有保护用户私钥的责任。

（8）审计与日志检查。为了安全起见，CA 对一些重要的操作应记入系统日志。在 CA 发生事故后，要根据系统日志做善后追踪处理——审计。CA 管理员要定期检查日志文件，尽早发现可能的隐患。

#### 5. 用户证书的吊销

在下列情形下，应当将用户证书吊销。

（1）一个用户证书到期。

（2）用户秘密密钥泄露。

（3）CA 的证书失窃。

（4）CA 不再给用户签发证书。

每一个 CA 必须维护一个证书吊销列表（Certificate Revocation List，CRL）。CRL 中列出所有已吊销证书的序列号和吊销日期。

### 3.3.2　X.509 证书标准

为了保障数字证书合理获取、撤出和验证过程，1988 年 ITU-T 发表了 X.509 标准。这是一个基于公开密钥和数字签名的标准，它的核心是数字证书格式和认证协议。X.509 作为

X.500 目录服务的一部分，定义了下列内容。

（1）X.500 目录向用户提供认证业务的一个框架。

（2）证书格式。

（3）基于公钥证书的认证协议。

### 1．X.509 数字证书结构

X.509 标准的核心是与用户有关的公开密钥证书。其 V3 的结构如下。

- Certificate（证书）
  - Version（版本）
  - Serial Number（序列号）
  - Algorithm ID（算法标识）
  - Issuer（颁发者）
  - Validity（有效期）
    - Not Before（起始日期）
    - Not After（终止日期）
  - Subject（使用者）
  - Subject Public Key Info（使用者公钥信息）
    - Public Key Algorithm（公钥算法）
    - Subject Public Key（公钥）
  - Issuer Unique Identifier（Optional）（颁发者的唯一标识）
  - Subject Unique Identifier（Optional）（使用者的唯一标识）
  - Extensions（Optional）（扩展）
    - …
- Certificate Signature Algorithm（证书签名算法）
- Certificate Signature（证书签名）

X.509 标准使用了下面的描述进行证书定义：

$$CA \ll A \gg = CA\{V, SN, AI, CA, T_A, A, A_P\}$$

其中：

Y<<X>>表示证书发放机构 Y 向用户 X 发放的证书。

Y{I}表示 I 链接上 Y 对 I 的哈希值签名。

### 2．证书目录

证书产生之后，必须以一定的方式存储和发布，以便于使用。X.509 标准的公开密钥证书由 CA 或用户放在 X.500 目录下进行集中存储和管理，并由一个可信赖的证书授权系统 CA 确认。在证书目录中，不仅存储和管理用户证书，还存储用户的相关信息（如电子邮件地址、电话号码等）。由于证书的非保密性，证书目录也是非保密的。

在标准化方面，目前证书目录广泛使用 X.500 标准。X.500 标准目录不仅可以对证书进

行集中管理，还可以管理用户的相关信息，从而构成一个用户信息源。

为了便于实际应用，在 Internet 环境下更多使用的是 X.500 标准的简化和改进版本——轻型目录访问协议（Lightweight Directory Access Protocol，LDAP）。

**3．X.509 证书的层次结构**

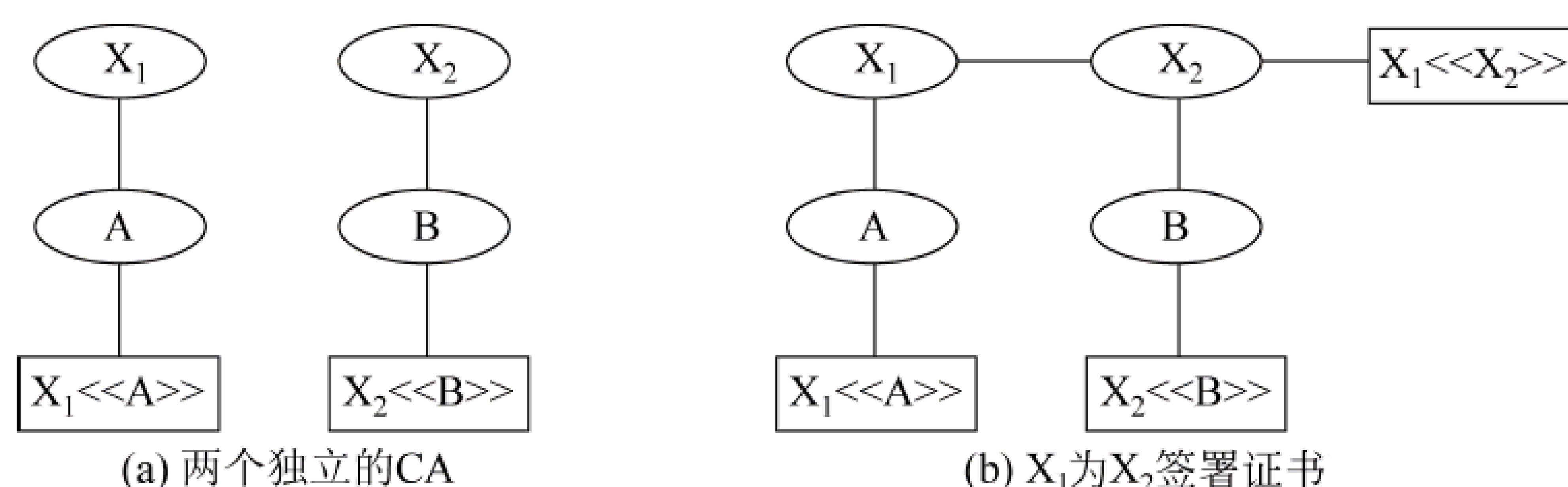
X.500 目录的作用是存放用户的公钥证书。由于证书不能伪造，它们可以不需要特别的保护就可以放在目录里。现在的问题是，是不是所有的证书都要由同一个 CA 签署？

一般说来，当用户数目较多时，仅由一个 CA 为所有用户签署是不现实的。因为这样需要两个条件。

（1）CA 必须取得所有用户的信任。

（2）每一个用户必须以绝对可靠的方式通过复制获得 CA 的公钥来证实。

显然，当用户数目较多时，应当由多个 CA 分头为不同的用户签署证书。但是，简单地由多个 CA 为不同的用户签署证书，也有一些问题。例如，$X_1$ 为 A 签署了一个证书，$X_2$ 为 B 签署了一个证书。那么，A 不能阅读 B 的证书，也不能证实 B 的证书；同样，B 不能阅读 A 的证书，也不能证实 A 的证书。这一目录结构如图 3.11（a）所示。



(a) 两个独立的CA　(b) $X_1$为$X_2$签署证书

图 3.11　多 CA 结构

观察图 3.11（b），情况就不同了。

（1）A 可以从此目录中获得 $X_1$ 签署的 $X_2$ 的证书。由于 A 确切知道 $X_1$ 的公钥，从 $X_1$ 的证书中就可以获得 $X_2$ 的公钥，并利用 $X_1$ 来证实。

（2）进一步 A 能获得 $X_2$ 签署的 B 的证书，并可以用已经获得的 $X_2$ 的公钥来证实 B 的数字签名，安全地获得 B 的公钥。

这样，就形成了证书链。其中，A 获得 B 的公钥的证书链，在 X.509 中的表示为

$$X_1<<X_2>>X_2<<B>>$$

同理，B 通过反向链也可以获得 A 的公钥，其结构表示为

$$X_2<<X_1>>X_1<<A>>$$

这样证书链形成一个层次结构。X.509 建议所有的 CA 证书须由 CA 放在目录中，并且要采用层次结构。图 3.12 为 X.509 层次结构的一个例子，其内部节点表示 CA，叶节点表示用户。用户可以从目录中沿着一条证书路径获得另一个节点的证书和公钥。例如，A 获取 B 证书的证书路径为

$$X<<W>>W<<V>>V<<Y>>Y<<Z>>Z<<B>>$$

A 取得这些证书后，就能解密其证书路径，获得一个可信的 B 的公钥。

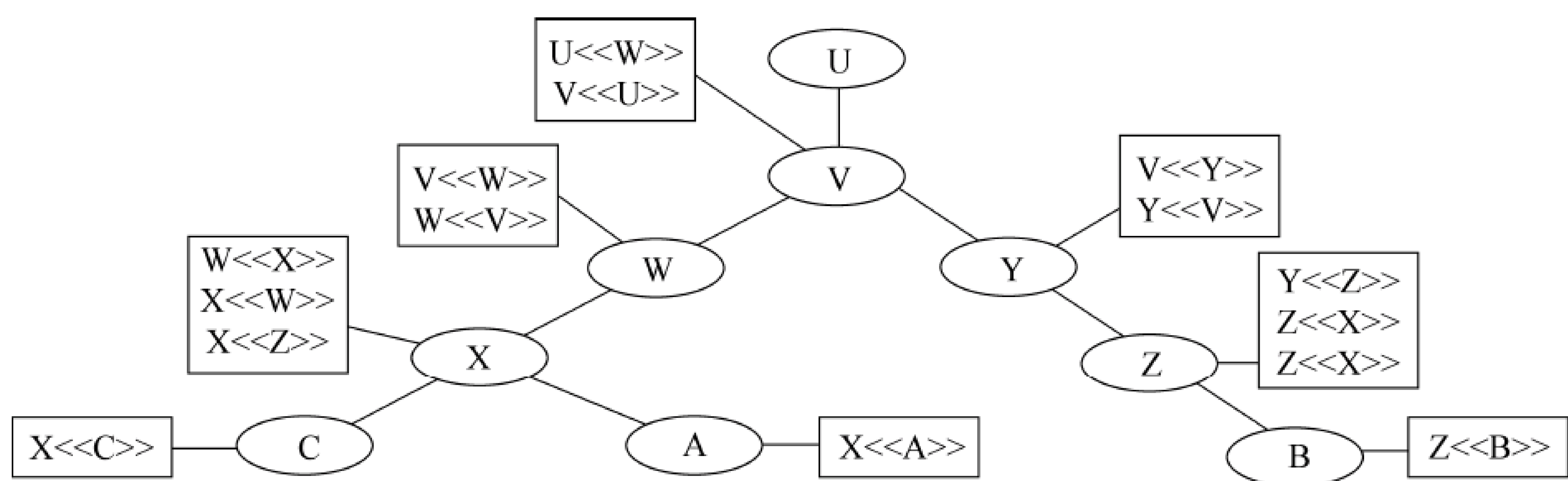



图 3.12　X.509 层次结构的一个例子

## 4．X.509 验证过程

图 3.13 为 X.509 建议的 3 种验证过程：一次验证过程、二次验证过程和三次验证过程。

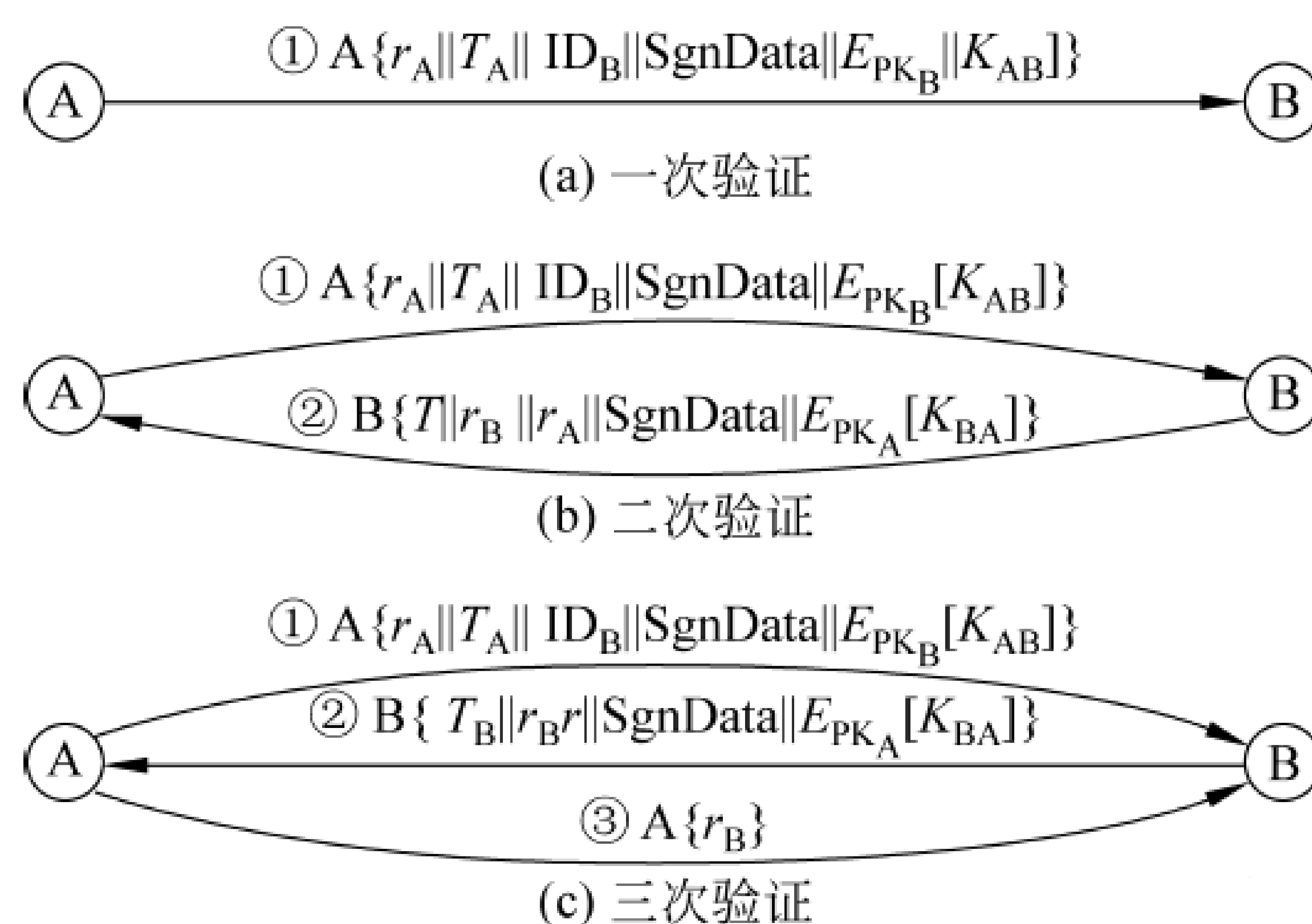



图 3.13　X.509 的 3 种验证过程

这 3 种验证过程都是使用公钥签名技术，并假定通信双方都认可目录服务器获得对方的公钥证书，或对方最初发来的报文中包括公钥证书（即双方都知道对方的公钥）。

1）一次验证

一次验证也称为单向验证。被验证者 A 产生报文供验证者 B 验证。包括如下内容。

（1）$ID_B$：B 的身份。

（2）$T_A$：时间戳，以保证报文的新鲜性。其中可以包括报文产生的时间（可选）和截止时间，以处理报文传送过程中可能出现的时延。

（3）$r_A$：一次性随机数，防止重放。在报文未到截止时间前是唯一的，以拒绝具有相同 $r_A$ 的其他报文。

如果仅仅为了验证，可以上述报文作为凭证；否则，还应包括下列内容。

（1）A 用自己的公钥签署的数字签名 SgnData，以保证信息的真实性和完整性。

（2）由 B 的公钥加密的双方会话密钥 $K_{AB}$。

2）二次验证

二次验证也称为双方验证，即 A 不仅要向 B 发送验证凭证信息，B 也要通过应答以证明以下几点。

（1）$ID_B$ 的身份。

（2）应答是由 B 发出的。

（3）应答的接收者是 A。

（4）应答报文是完整的和新鲜的。

3）三次验证

三次验证是在二次验证完成之后，A 再将 B 发来的一次性随机数签名后发往 B。这样通过检查一次性随机数就可以得知是否有重放，而不需要检查时间戳。这种方法主要用在通信双方无法建立时钟同步的情形。

### 3.3.3　公开密钥基础设施

#### 1．PKI 及其职能

公开密钥基础设施（Public Key Infrastructure，PKI）是 20 世纪 80 年代在公开密钥理论和技术的基础上发展起来的为电子商务提供综合、安全基础平台的技术和规范。它的核心是对信任关系的管理。通过第三方信任，为所有网络应用透明地提供加密和数字签名等密码服务所必需的密钥和证书管理，从而达到保证网上传递数据的安全、真实、完整和不可抵赖的目的。PKI 的基础技术包括加密、数字签名、数据完整性机制和双重数字签名等。利用 PKI 可以方便地建立和维护一个可信的网络计算环境，建立一种信任机制，使人们在这个无法相互见面的环境中，能够确认对方的身份和信息，从而为电子支付、网上交易、网上购物和网上教育等提供可靠的安全保障。

PKI 系统的建立着眼于用户使用证书及相关服务的便利性以及用户身份认证的可靠性。具体职能如下。

（1）制定完整的证书管理政策。

（2）建立高可信度的 CA 中心。

（3）负责用户属性管理、用户身份隐私的保护和证书作废列表的管理。

（4）为用户提供证书和 CRL 有关服务的管理。

（5）建立安全和相应的法规，建立责任划分并完善责任政策。

因此，PKI 是一个使用公钥和密码技术实施并提供安全服务的、具有普适性的安全基础设施的总称，并不特指某一密码设备及其管理设备。可以说，它是生成、管理、存储、颁发和撤销基于公开密码的公钥证书所需要的硬件、软件、人员、策略和规程的综合。

#### 2．PKI 的组成

通常 CA 分成不同的一些层次。一个典型 PKI 体系结构如图 3.14 所示。其中，PAA 为政策批准机构，PCA 为政策认证机构，ORA（Online Registration Authority）为在线注册机构。它们的区别在于政策权限不同：下层的证书要由上层颁发。

1）政策批准机构

政策批准机构（PAA）是一个 PKI 系统方针的制定者，它建立整个 PKI 体系的安全策略，批准本 PAA 下属的 PCA 的政策，为下属 PCA 签发证书，并负有监控各 PCA 行为的责任。

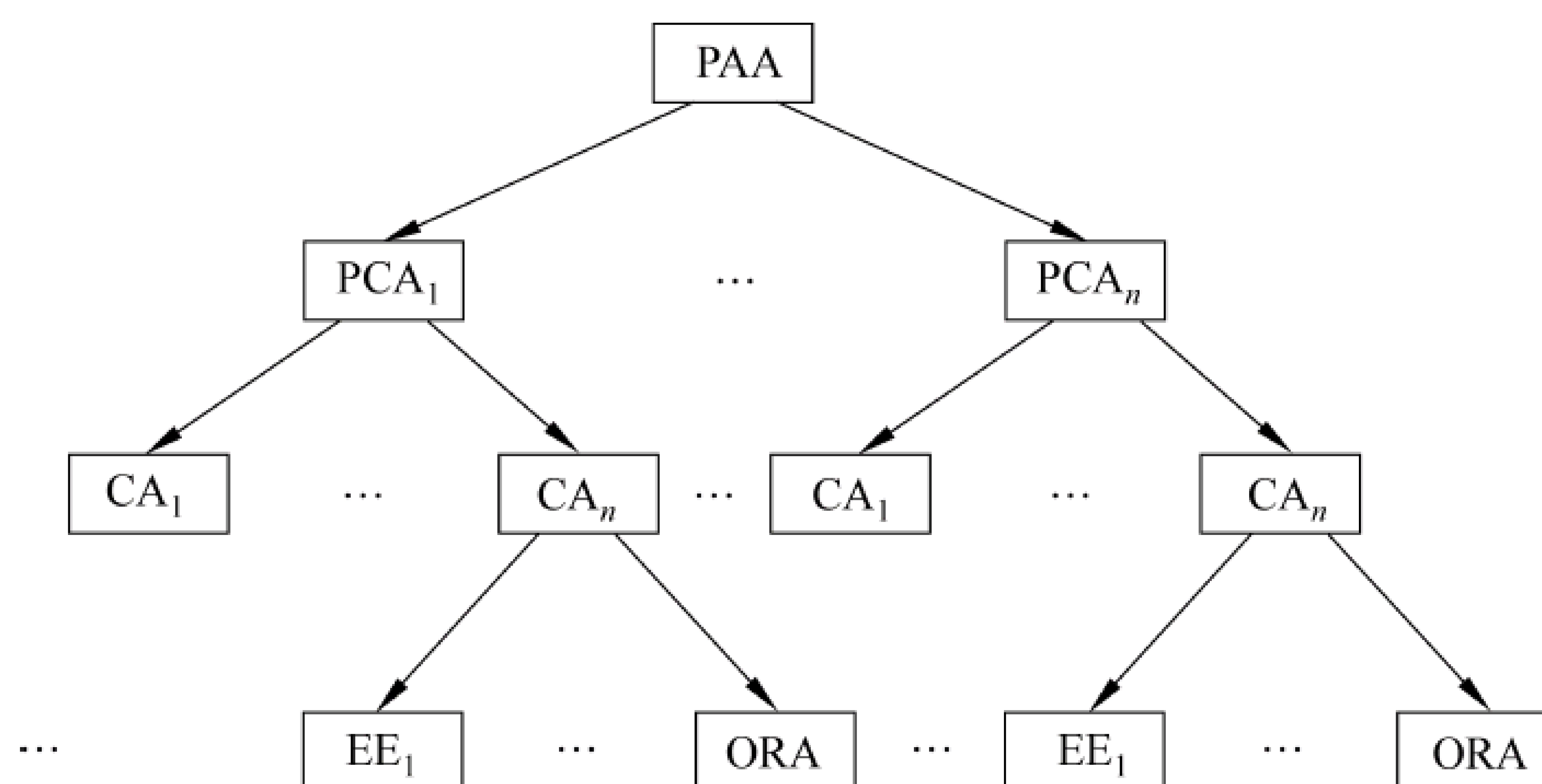


图 3.14　典型的 PKI 体系结构

2）政策认证机构

PCA 指定自身的具体政策。这些政策可以是其上级 PAA 政策的扩充或细化（包括本 PCA 范围内密钥的产生、密钥的长度、证书的有效期规定以及 CRL——被吊销的证书列表的管理），并为下属 CA 签发公钥证书。

3）认证机构

CA 具有有限政策制定权限，它在上级 PCA 政策范围内，进行具体的用户公钥证书的签发、生成和发布以及 CRL 的生成和发布。

4）在线注册机构

ORA 进行证书申请者的身份认证，向 CA 提交证书申请，验证接收 CA 签发的证书，并将证书发放给申请者。有时还协助进行证书的制作。

## 实验 8　证书制作及 CA 系统配置

### 1. 实验目的

（1）深入理解 PKI 系统的工作原理。
（2）掌握在一种系统中配置 CA 系统的方法。
（3）掌握证书的申请及制作方法。
（4）体会 SSL 的作用。

### 2. 实验内容

（1）选择一个系统，进行 CA 系统的配置。
（2）为服务器和浏览器之间的安全通信进行设置。
（3）为某些用户生成证书。
（4）测试上述通信的安全性。

### 3. 建议环境

（1）利用在 Windows 2000 Server 中附加的认证服务器，进行 Windows 2000 PKI 系统的

配置。

（2）利用 Linux 平台上的 SSL X.509 或 FHS 进行实验。

#### 4. 实验准备

（1）收集资料，设计在实验用系统中进行 CA 系统配置的步骤。

（2）设计在实验用系统中进行安全通信配置的步骤。

（3）设计为用户生成证书的方法和步骤。

（4）设计对上述系统配置进行通信安全测试的方法和步骤。

#### 5. 推荐的分析讨论内容

（1）你知道有哪些证书标准？

（2）你知道有哪些通信安全协议？试进行比较。

（3）其他发现或想到的问题。

## 3.4 信息系统访问授权

一个信息系统当然不允许非法用户访问，即使是合法用户，也不是就可以访问系统的所有资源或者对系统的某一资源进行为所欲为的访问操作。系统访问控制就是基于这种考虑的安全机制。

访问控制分为系统访问控制和网络访问控制。

系统访问控制是从系统资源安全保护的角度对访问进行授权控制。它从访问的角度将系统对象分为主体（subject）和客体（object）两类。主体也称为访问发起者，主要指用户、用户组、进程以及服务等；客体也称为资源，主要指文件、目录和计算机等。授权就是赋予主体一定的权限（修改、查看等），赋予客体一定的访问属性（如读、写、添加、执行、发起连接等），同时在主体与客体之间建立一套安全访问规则，通过对客体的读出、写入、修改、删除、运行等的管理，确保主体对客体的访问是经过授权的，同时要拒绝非授权的访问，以保证信息的机密性、完整性和可用性。

### 3.4.1 访问控制的二元关系描述

访问控制用一个二元组（控制对象，访问类型）来表示。其中的控制对象表示系统中一切需要进行访问控制的资源，访问类型是指对于相应的受控对象的访问控制，如读取、修改、删除等。

访问控制二元组有许多描述形式。下面介绍几种常用的形式。

#### 1. 访问控制矩阵

访问控制矩阵也称为访问许可矩阵，它用行表示客体，列表示主体，在行和列的交叉点上设定访问权限。表 3.1 是一个访问控制矩阵的例子。表中，一个文件的 Own 权限的含义是可以授予（authorize）或者撤销（revoke）其他用户对该文件的访问控制权限。例如，

张三对 File1 具有 Own 权限，所以张三可以授予或撤销李四和王五对 File1 的读（R）和写（W）权限。

**表 3.1　一个访问控制矩阵的例子**

| 主体 | 客体 | | | |
|---|---|---|---|---|
| | File1 | File2 | File3 | File4 |
| 张三 | Own,R,W | | Own,R,W | |
| 李四 | R | Own,R,W | W | R |
| 王五 | R,W | R | | Own,R,W |

### 2. 授权关系表

授权关系表（authorization relations）描述了主体和客体之间各种授权关系的组合。表 3.2 为表 3.1 的授权关系表。

**表 3.2　授权关系表的一个例子**

| 主体 | 访问权限 | 客体 |
|---|---|---|
| 张三 | Own | File1 |
| 张三 | R | File1 |
| 张三 | W | File1 |
| 张三 | Own | File3 |
| 张三 | R | File3 |
| 张三 | W | File3 |
| 李四 | R | File1 |
| 李四 | Own | File2 |
| 李四 | R | File2 |
| 李四 | W | File2 |
| 李四 | W | File3 |
| 李四 | R | File4 |
| 王五 | R | File1 |
| 王五 | W | File1 |
| 王五 | R | File2 |
| 王五 | Own | File4 |
| 王五 | R | File4 |
| 王五 | W | File4 |

授权关系表便于使用关系数据库进行存储。只要按照客体进行排序，就得到了与访问能力表相当的二维表；按照主体进行排序，就得到了与访问控制表相当的二维表。

例如，当用户或应用程序试图访问一个文件时，首先需要通过系统调用打开文件。在打开文件之前，访问控制机制被调用。访问控制机制利用访问控制表、访问能力表或访问控制矩阵等，检查用户的访问权限，如果在用户的访问权限内，则可以继续打开文件；如

果用户超出授权权限，则访问被拒绝，产生错误信息并退出。

**3. 访问能力表**

能力（capability）也称为权能，是受一定机制保护的客体标志，标记了某一主体对客体的访问权限：某一主体对某一客体有无访问能力，表示了该主体能不能访问那个客体；而具有什么样的能力，表示能对那个客体进行一些什么样的访问。它也是一种基于行的自主访问控制策略。图 3.15 是表 3.1 所示的访问控制矩阵的访问能力表表示。

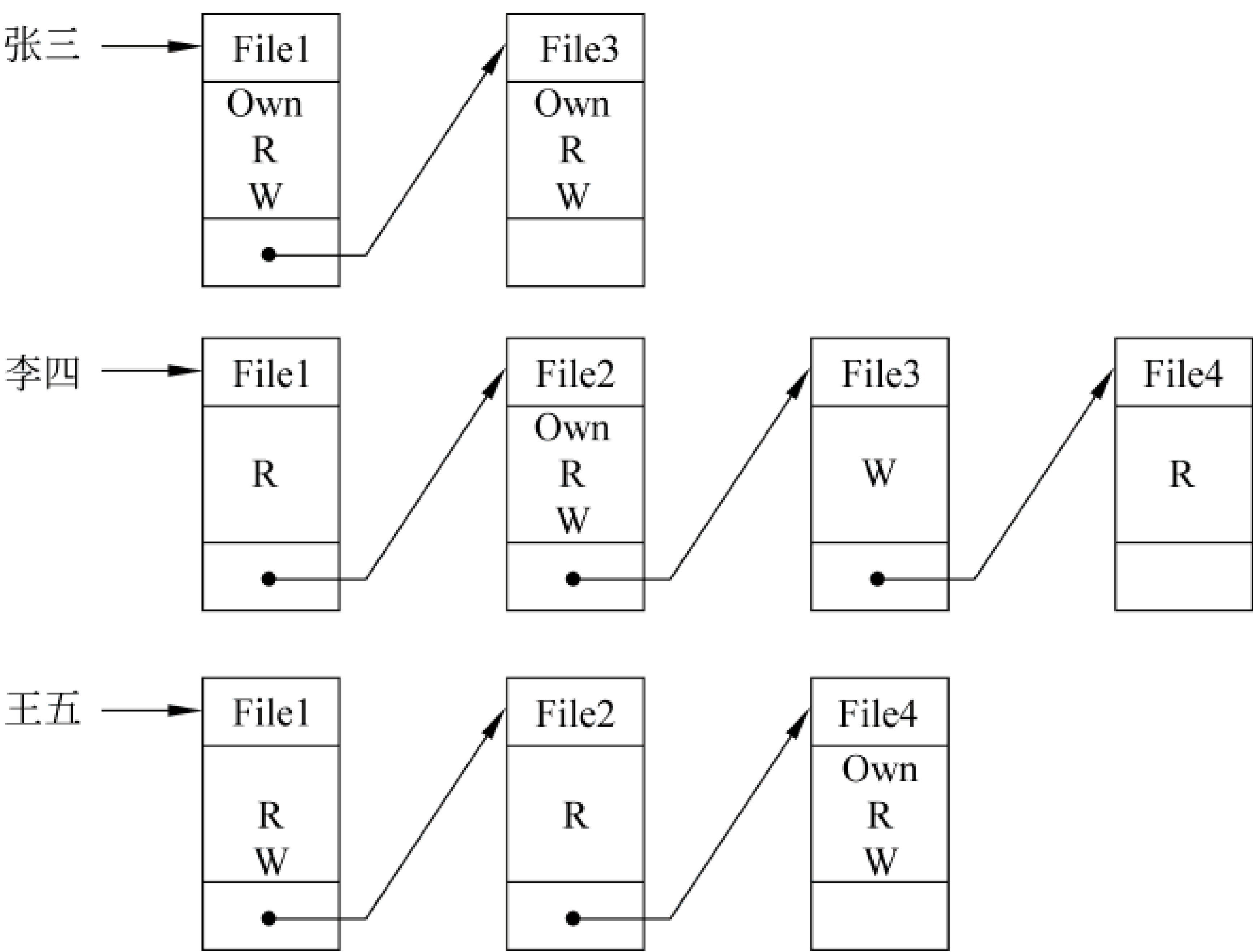



图 3.15　访问能力表的例子

访问能力表允许在进程运行期间动态地发放、回收、删除或增加某些权力，执行速度比较快，还可以定义一些系统事先不知道的访问类型。此外，访问能力表着眼于某一主体的访问权限，从主体出发描述控制信息，很容易获得一个主体所被授权可以访问的客体及其权限，但要从客体出发获得哪些主体可以访问它就困难了。目前使用访问能力表实现的自主访问控制系统已经不多。

**4. 访问控制表**

访问控制表（Access Control List，ACL）与访问能力表正好相反，是从客体出发描述控制信息，可以用来对某一资源指定任意一个用户的访问权限。这种方式给每个客体建立一个 ACL（访问控制表），记录该客体可以被哪些主体访问以及访问的形式。它是一种基于列的自主访问控制策略。图 3.16 是表 3.1 的访问控制表表示。可以看出，每个 ACL 包括一个 ACL 头和零个或多个 ACE（访问控制项）。

ACL 的优点是可以很容易地查出对某一特定资源拥有访问权的所有用户，有效地实施授权管理，是目前采用得最多的一种实现形式。Windows NT/2000/XP 的资源（文件、设备、邮件槽、已命名的和未命名的管道、进程、线程、事件、互斥体、信号量、可等待定时器、访问令牌、窗口站、网络共享、服务、注册表和打印机等）访问就是采用这种方式。

ACL 适合按照对象进行访问的操作系统。但是使用 ACL 进行访问权限的管理，仅依靠

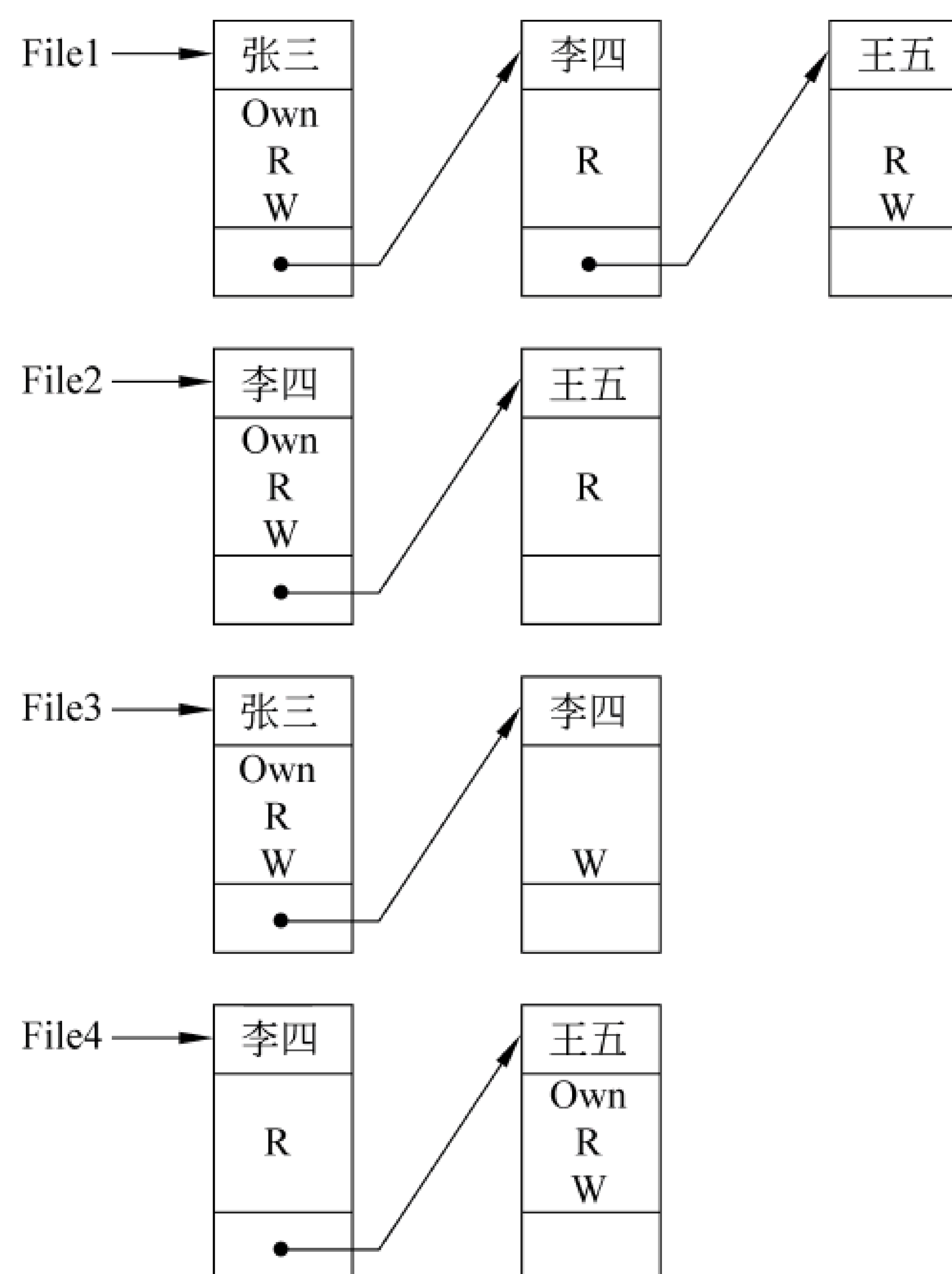



图 3.16　访问控制表的例子

单个主体非常麻烦。为此，通常将用户按组进行组织，用户也可以从用户组取得访问权限。在 UNIX 中，附在文件上的简单的 ACL 允许对用户、组和其他三类主体规定基本访问模式。

## 3.4.2　自主访问控制与强制访问控制

资源的所有者往往是资源的创建者。大多数操作系统支持资源所有权的概念，并且在决定访问控制策略时考虑资源所有权。基于所有权的访问控制可以有两种基本的策略：自主访问控制（Discretionary Access Control，DAC）和强制访问控制（Mandatory Access Control，MAC）。

### 1. 自主访问控制策略

自主访问控制是目前计算机系统中应用最广泛的一种策略，主流操作系统 Windows Server、UNIX 系统，以及防火墙（ACL）等都是采用自主型的访问控制策略。它的基本思想是，资源的所有者可以对资源的访问进行控制，任意规定谁可以访问其资源，自主地直接或间接地将权限传给（分发给）主体。例如，用户 A 对客体 O 具有访问权限，而 B 没有。当 A 将对 O 的访问权限传递给 B 后，B 就有了对 O 的访问权限。

口令（password）机制就是一种基于行的自主访问控制策略。它要求每个客体都相应地有一个口令。主体对客体进行访问前，必须向操作系统提供该客体的口令。采用这种机制的系统有 IBM 公司的 MVS 和 CDC 公司的 MOS 等。

DAC 的优点是应用灵活与可扩展，所以经常被用于商业系统。其缺点是，权限传递很容易造成漏洞，安全级别比较低，不太适合网络环境，主要用于单个主机。

通常 DAC 通过访问控制矩阵来限定哪些主体针对哪些客体可以执行什么操作。但是，

目前操作系统在实现自主访问控制时，不是利用整个访问控制矩阵，而是基于访问控制矩阵的行或列来表达访问控制信息。这样就可以非常灵活地对策略进行调整。

**2. 强制访问控制策略**

强制访问控制（MAC）也称为系统访问控制，它的基本思想是系统要“强制”主体服从访问控制政策：系统（系统管理员）给主体和客体分配了不同的安全属性，用户不能改变自身或任何客体的安全属性，即不允许单个用户确定访问权限，只有系统管理员才可以确定用户或用户组的访问权限。

MAC 主要用于多层次安全级别的系统（如军事系统）中。它预先将主体和客体进行分级，定义出一些安全等级（如高密级、机密级、秘密级、无密级等）并用对应的标签进行标识：对于主体称为许可级别和许可标签，对于客体称为安全级别和敏感性标签。用户必须遵守依据安全策略划分的安全级别的设定以及有关访问权限的设定。

由于主体有既定的许可级别，客体也有既定的安全级别，因此主体对客体能否执行特定的操作，取决于两者的安全属性之间的关系。例如，对于信息（文件）的访问，可以定义如下 4 种关系。

（1）下读（read down）：用户级别高于信息级别的读操作。

（2）上读（read up）：用户级别低于信息级别的读操作。

（3）下写（write down）：用户级别高于信息级别的写操作。

（4）上写（write up）：用户级别低于信息级别的写操作。

当用户提出访问请求时，系统对主体和客体的安全属性进行比较，来决定该主体是否可以对所请求的客体进行访问。当一个主体（进程）要访问客体，其许可标签必须满足下面的条件。

（1）主体若要对客体具有写访问的权限，则其许可级别必须被客体的安全级别支配。

（2）主体若要对客体具有读访问的权限，则其许可级别必须支配被客体的安全级别。

在典型的应用中，MAC 使用两种访问控制关系：上读/下写（用来保证数据完整性）和下读/上写（用来保证数据机密性）。下读/上写相当于在一个层次组织中，上级可以看下级的资料，而下级不能看上级的资料，但可以向上级写资料，图 3.17 为下读/上写的示意图。

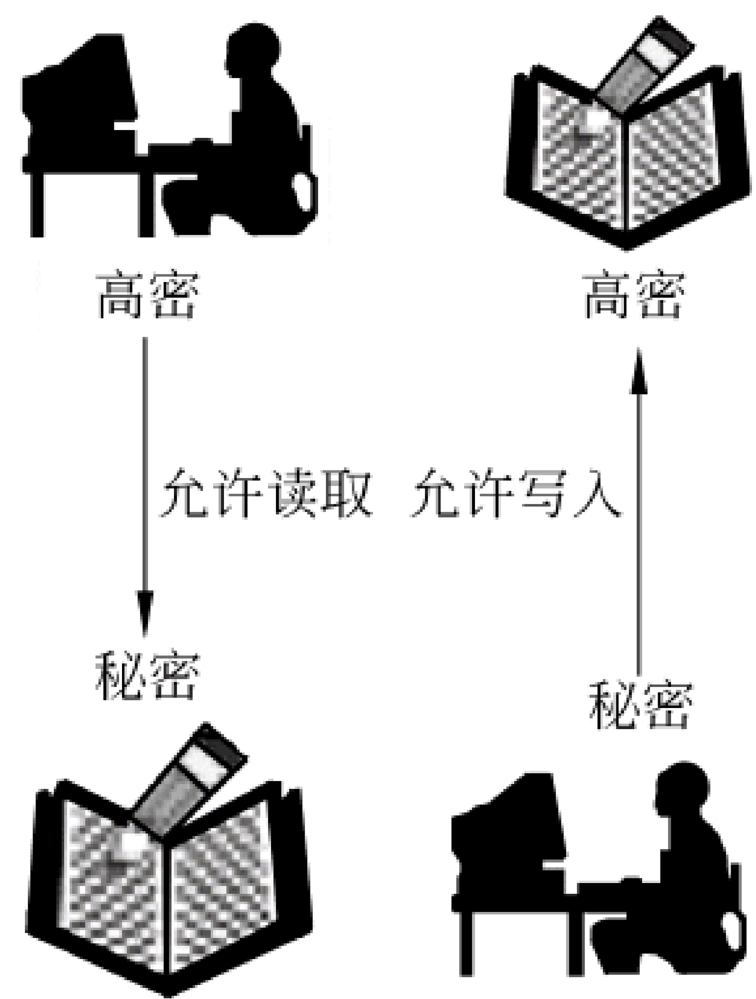


图 3.17　下读/上写示意图

MAC 比 DAC 具有更强的访问控制能力，但是实现的工作量大，管理不便，不够灵活。

强制访问控制和自主访问控制有时会结合使用。例如，系统可能首先执行强制访问控制来检查用户是否有权限访问一个文件组（这种保护是强制的，也就是说，这些策略不能被用户更改），然后再针对该组中的各个文件制定相关的访问控制表（自主访问控制策略）。

### 3.4.3 基于角色的访问控制策略

角色（role）是指一个组织或任务中的岗位、职位或分工。角色需要人去扮演。一般说来，一个角色并非只有一人扮演，如会计这个角色往往需要多个人；并且一个人可能会从事不同的角色。基于角色的访问控制（Role-Base Access Control，RBAC）就是基于这样一种考虑而提出的访问控制策略。由于角色比个体用户具有较大的稳定性，这种授权管理比针对个体的授权管理在可操作性和可管理性方面都要强得多。

如图 3.18 所示，角色实际上是在主体（用户）与客体之间引入的中间控制机制层。

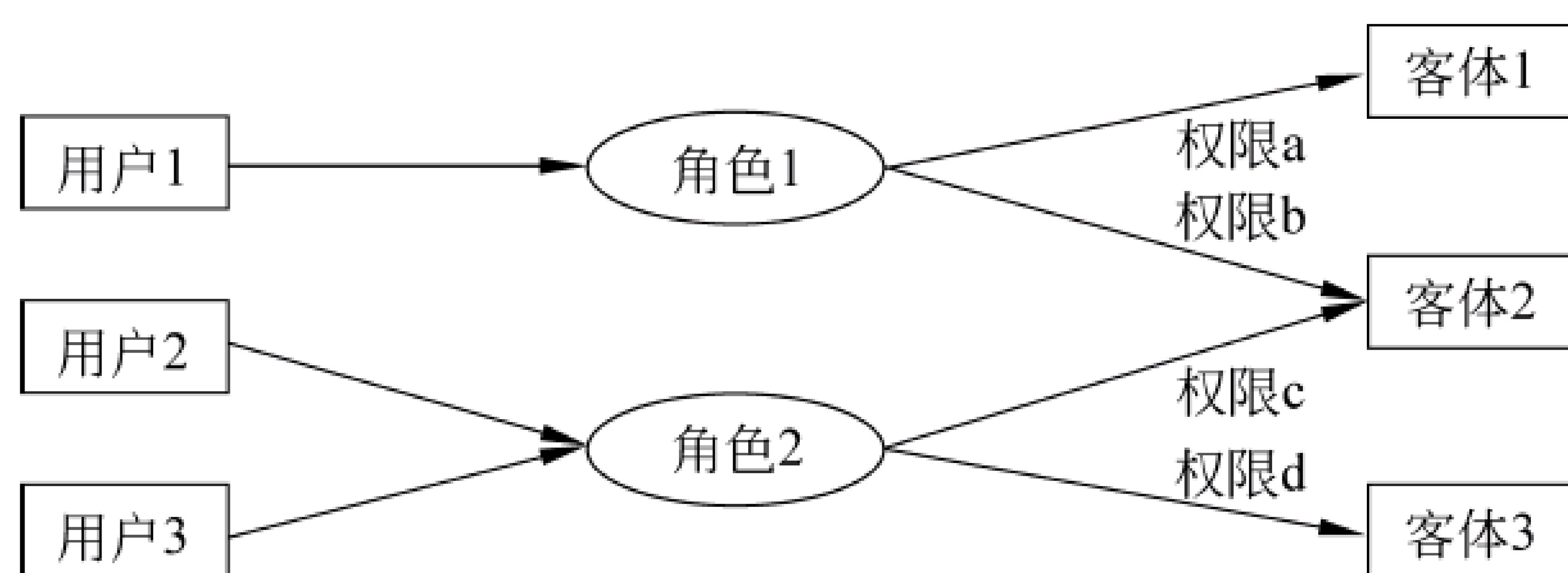



图 3.18　角色是在主体与客体之间引入的中间控制机制层

在 RBAC 系统中，要求明确地区分权限（authority）和职责（responsibility）或区分操作与管理，使两者互相制约。

**例 3.1**　一位科长可以对所在的科里的成员发号施令，而并不能对其他科的科员发号施令。因为从权力上看，他是科长，而从职责上来说，他只是某一个科的科长，并非所有科的科长。与之相仿，对于一个具有高密级（0 级）许可级的用户来说，并不可以访问所有安全级别为 0 级的资源。因为有些资源不在他的职责范围内。

**例 3.2**　一个可以访问某个资源集合的用户，并不能进行该资源集合的访问授权。因为他没有这个权限。

**例 3.3**　一位安全主管有权进行授权分配，但不能同时具有访问数据资源的权力。

由于实现了权限与职责的逻辑分离，基于角色的策略极大地方便了权限管理。例如，如果一个用户的职位发生变化，只要将用户当前的角色去掉，加入代表新职务或新任务的角色即可。基于角色的访问控制方法还可以很好地描述角色层次关系，实现最少权限原则和职责分离的原则，非常适合在数据库应用层的访问控制。因为在应用层，角色的概念比较明显。

角色由系统管理员定义，角色成员的增减也只能由系统管理员执行，只有系统管理员才有权定义和分配角色，并且授权规则是强加给用户的，用户只能被动地接受，不能自主地决定。但是，角色的控制比较灵活，根据需要可以将某些角色配置得接近 DAC，而让某些角色接近 MAC。

## 实验 9　用户账户管理与访问权限设置

### 1. 实验目的

（1）掌握在一个系统中进行用户账户管理的方法。

（2）掌握在一个系统中进行访问权限设置的方法。

**2. 实验内容**

（1）在一个系统中进行用户账户管理（若是在 Linux 系统中，需考虑添加批量用户）和安全设置。

（2）在一个系统中进行访问权限设置（若是在 Windows 2000 以上的系统中，需基于 NTFS 进行设置）。

**3. 建议环境**

在一种操作系统环境（如 Linux 或 Windows）下设置账号和访问权限，进行用户管理。

**4. 实验示范——Windows 中账户和权限的设置**

Windows 2000/NT 拥有强大的用户和组权限管理功能，在保护系统安全方面有着独特的应用。通过为不同用户分配相应权限，可以限制其对系统重要文件或目录的访问，并以此达到保护系统不受到病毒和黑客侵犯的功能。

1）Windows 中的账户设置

（1）单击"我的电脑"→"控制面板"，打开"控制面板"对话框，选择"用户账户"，如图 3.19 所示。




图 3.19 "控制面板"中的"用户账户"

（2）双击"用户账户"，进入"用户账户"窗口，可以看到在"挑选一项任务…"中有一些选项，如图 3.20 所示。




图 3.20 “用户账户”窗口

（3）选择“创建一个新账户”，进入“用户账户”窗口中的“为新账户起名”页面，可以为新账户起名，如图 3.21 所示。

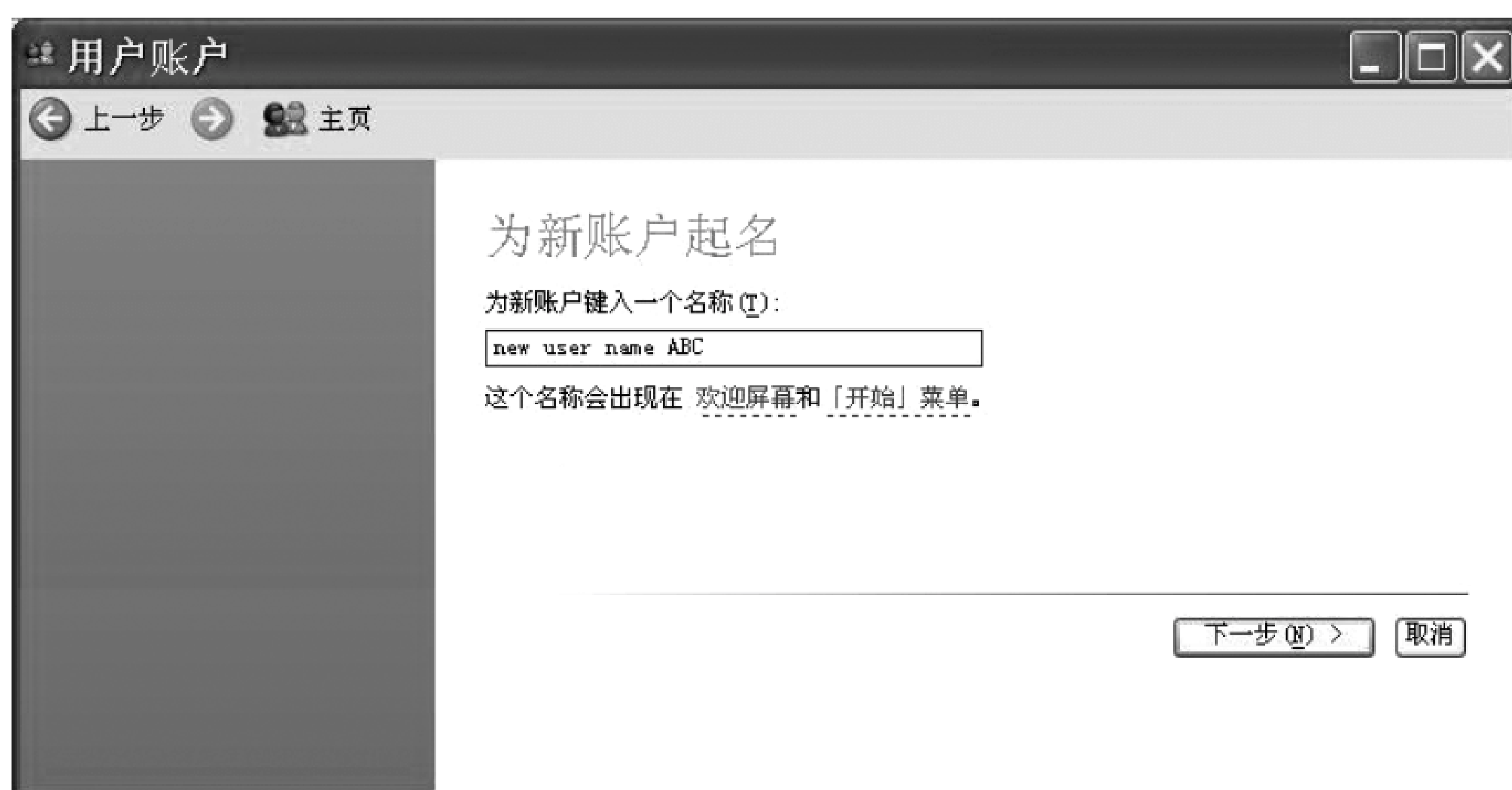



图 3.21 为新账户起名

（4）为新用户输入一个名称后，单击“下一步”按钮，打开“挑选一个账户类型”页面，可以为新用户选择一个账户类型。图 3.22 为选择“计算机管理员”选项时的页面。

这类账户的权力如下。

① 创建、更改和删除账户。

② 进行系统范围的更改。

③ 安装程序并访问所有文件。

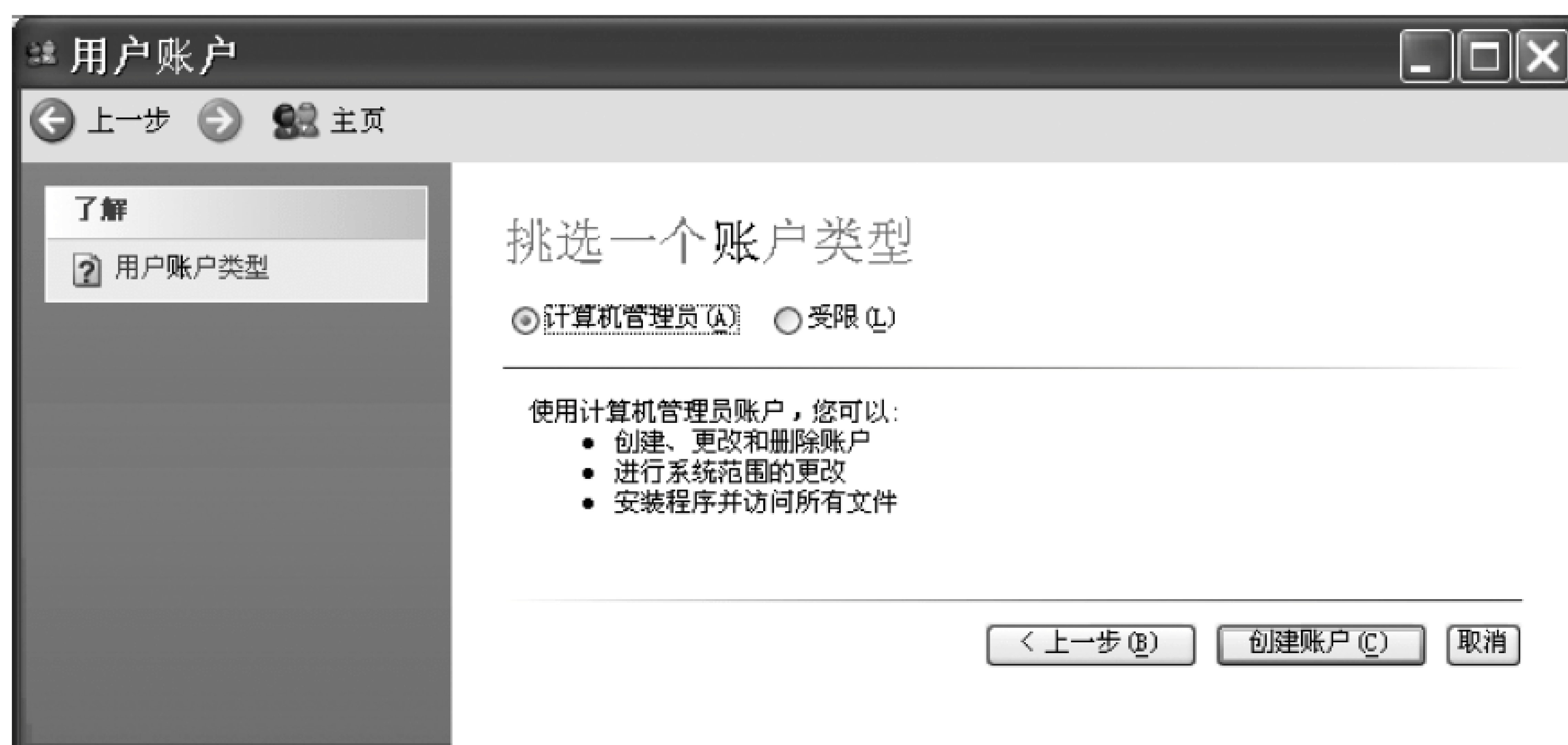



图 3.22　挑选一个账户类型——计算机管理员

图 3.23 为选择“受限”账户类型时的页面。

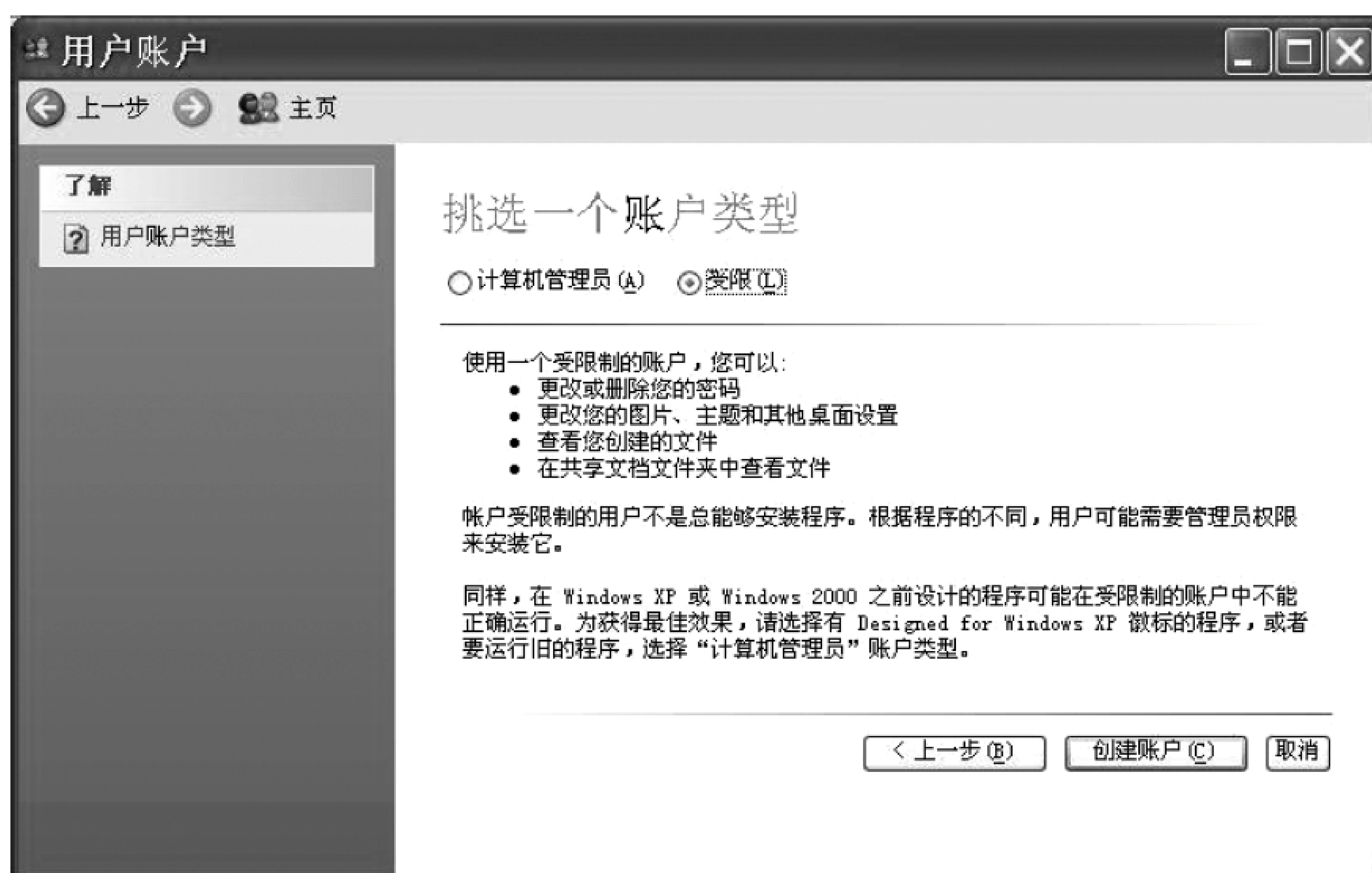



图 3.23　挑选一个账户类型——受限

这类账户的权力如下。

① 更改或删除自己的密码。

② 更改自己的图片、主题和其他桌面设置。

③ 查看自己创建的文件。

④ 在共享文档文件夹中查看文件。

（5）用户类型选择后，单击“创建账户”按钮，系统进入如图 3.24 所示的页面，提示对新建账户可以进行的操作选项。

（6）选择“创建密码”选项，进入如图 3.25 所示的页面，在“输入一个新密码”文本框中输入密码。

这样，以后再次启动系统时，就会要求先输入账号和密码。不同的账户也可以把自己的私人文档保存到“我的文档”文件夹中，把该文件夹设置为专用。




图 3.24　选择对新建账户的操作

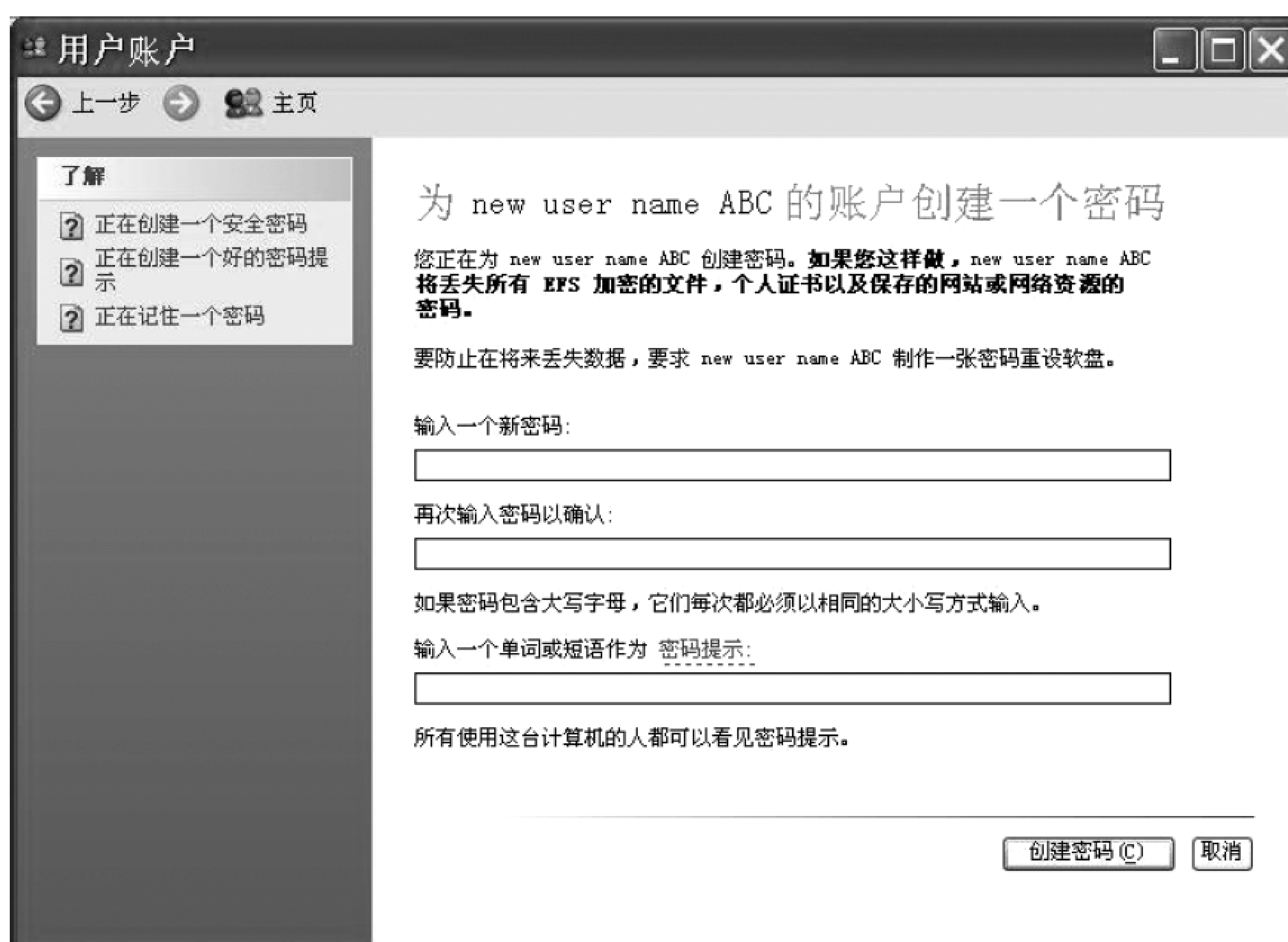



图 3.25　为新账户创建密码

2）Windows 2000/NT 中的组策略

在 Windows 2000/NT 中，用户被分成许多组，组和组之间都有不同的权限。当然，一个组的用户和用户之间也可以有不同的权限。下面是一些常用的组。

（1）Administrators（管理员组）。

管理员可以执行操作系统所支持的所有功能。Windows 2000/NT 默认安全设置不限制管理员对任何注册表或文件系统对象的访问。只有受信任的人员才可以成为该组成员。

（2）Power Users（高级用户组）。

在权限设置中，这个组的权限是仅次于 Administrators 组的。Power Users 可以执行除了

为 Administrators 组保留的任务以外的其他任何操作系统任务。分配给 Power Users 组的默认权限允许 Power Users 组的成员修改整个计算机的设置，但 Power Users 不具有将自己添加到 Administrators 组的权限。

（3）Users（普通用户组）。

Users 组提供了一个最安全的程序运行环境，默认安全设置旨在禁止该组的成员危及操作系统和已安装程序的完整性。系统对这个组赋予的权限如下。

① 该组的用户可以运行经过验证的应用程序，但不可以运行大多数旧版应用程序。

② Users 可以关闭工作站，但不能关闭服务器。

③ Users 可以创建本地组，但只能修改自己创建的本地组。

④ Users 不能修改系统注册表设置、操作系统文件或程序文件，不允许该组成员修改操作系统的设置或用户资料。

（4）Guests（来宾组）。

Guests 与普通 Users 的成员有同等访问权，但来宾账户的限制更多。

（5）Everyone（所有的用户）。

计算机上的所有用户都属于这个组。

实际上，还有一个组也很常见，它拥有和 Administrators 一样甚至比其更高的权限，但是这个组不允许任何用户的加入，在查看用户组的时候，它也不会被显示出来，它就是 System 组。

3）Windows 2000 的 NTFS 系统

新技术文件系统（New Technology File System，NTFS）是 Microsoft 公司为了弥补文件分配表（File Allocation Table，FAT）系统的一些不足而推出的一项技术，其最大的改进就是容错性和安全性能。

基于 NTFS 卷进行访问权限设置非常简单。右击一个 NTFS 卷或 NTFS 卷下的一个目录，在快捷菜单中选择“属性”→“安全”选项就可以对一个卷或者一个卷下面的目录进行权限设置。这时会看到以下 7 种权限。

（1）“完全控制”就是对此卷或目录拥有不受限制的完全访问，像 Administrators 在所有组中的地位一样。选中了“完全控制”，下面的 5 项属性将被自动选中。

（2）“修改”则像 Power Users，选中了“修改”，下面的 4 项属性将被自动选中。当下面的任何一项没有被选中时，“修改”条件将不再成立。

（3）“读取和运行”就是允许读取和运行在这个卷或目录下的任何文件。“列出文件夹目录”和“读取”是“读取和运行”的必要条件。

（4）“列出文件夹目录”是指只能浏览该卷或目录下的子目录，不能读取，也不能运行。

（5）“读取”是指能够读取该卷或目录下的数据。

（6）“写入”就是能往该卷或目录下写入数据。

（7）“特别”则是对以上 6 种权限进行细分。

### 5. 实验准备

（1）设计在一个系统中进行用户账户管理的步骤。
（2）设计在一个系统中进行访问权限设置的步骤。

### 6. 推荐的分析讨论内容

（1）在一个共用系统中，应当根据管理权限将系统的访问权限分成不同等级。请分析当每个人都具有自己的非私密性文件时，为保证系统的安全，比他的权限高和权限低的人分别应当在读、写和执行 3 种访问权限方面有何限制？
（2）其他发现或想到的问题。

# 习　题

## 一、选择题

1. 用数字办法确认、鉴定、认证网络上参与信息交流者或服务器的身份是指________。
   A. 接入控制　　B. 数字认证　　C. 数字签名　　D. 防火墙
2. 身份鉴别是安全服务中的重要一环，以下关于身份鉴别的叙述中不正确的是________。
   A. 身份鉴别是授权控制的基础
   B. 身份鉴别一般不用提供双向的认证
   C. 身份鉴别目前一般采用基于对称密钥加密或公开密钥加密的方法
   D. 数字签名机制是实现身份鉴别的重要机制
3. 以下关于CA认证中心说法正确的是________。
   A. CA认证是使用对称密钥机制的认证方法
   B. CA认证中心只负责签名，不负责证书的产生
   C. CA认证中心负责证书的颁发和管理，并依靠证书证明一个用户的身份
   D. CA认证中心不用保持中立，可以随便找一个用户来作为CA认证中心
4. Kerberos的设计目标不包括________。
   A. 认证　　B. 授权　　C. 记账　　D. 审计
5. 访问控制是指确定________以及实施访问权限的过程。
   A. 用户权限　　B. 可赋予哪些主体访问权利
   C. 可被用户访问的资源　　D. 系统是否遭受入侵
6. 以下各项中对访问控制影响不大的是________。
   A. 主体身份　　B. 客体身份　　C. 访问类型　　D. 主体与客体的类型
7. 为了简化管理，通常对访问者________，以避免访问控制表过于庞大。
   A. 分类组织成组
   B. 严格限制数量
   C. 按访问时间排序，删除长期没有访问的用户

D. 不作任何限制

8. PKI支持的服务不包括________。

A. 非对称密钥技术及证书管理　　B. 目录服务

C. 对称密钥的产生和分发　　D. 访问控制服务

9. PKI的主要组成不包括________。

A. 证书授权CA　　B. SSL　　C. 注册授权RA　　D. 证书存储库CR

10. PKI管理对象不包括________。

A. ID和口令　　B. 证书　　C. 密钥　　D. 证书撤销

11. 下面不属于PKI组成部分的是________。

A. 证书主体　　B. 使用证书的应用和系统

C. 证书权威机构　　D. AS

12. PKI能够执行的功能是________和________。

A. 鉴别计算机消息的始发者　　B. 确认计算机的物理位置

C. 保守消息的机密　　D. 确认用户具有的安全性特权

13. PKI的主要理论基础是________。

A. 对称密码算法　　B. 公钥密码算法　　C. 量子密码　　D. 摘要算法

## 二、填空题

1. ________是验证信息发送者是真的，而不是冒充的，包括________、________等的认证和识别。

2. ________的目的是为了限制访问主体对访问客体的________。

3. ________是PKI的核心元素，________是PKI的核心执行者。

## 三、问答题

1. 简述生物特征身份认证的发展趋势。

2. 简述口令可能会遭受哪些攻击。

3. 假定只允许使用26个字母构造口令，在下列情况下各可以构造出多少条口令？

（1）口令最多可以使用 $n$ 个字符，$n=4$，6，8，不区分大小写。

（2）口令最多可以使用 $n$ 个字符，$n=4$，6，8，区分大小写。

4. 编写一个口令生成程序。程序以长度 $s$（可以取 $s=8$，16，32，64）的随机二进制种子作为输入。

（1）让多名用户使用该程序生成口令，记录有多少人选择了相同的事件。

（2）生成一个口令并加密。然后让人通过尝试随机数种子的所有值进行口令攻击。事先要给定一个猜测次数的期望值。

5. 简述口令可能会遭受哪些攻击。

6. 比较动态口令的3种实现方式。

7. 比较静态口令与动态口令。

8. 常用动态令牌有哪几种？

（1）短信密码。

（2）手机令牌。

（3）硬件令牌。
（4）软件令牌。
9. 在身份验证中，可能会遇到重放攻击。重放具有如下几种形式。
（1）简单的重放。攻击者简单地复制信息，经过一段时间后，再重放原来的信息。
（2）重放不能被检测到。这时，原始的信息不能到达，只有重放信息到达目的地。
（3）没有定义的重放返回。发送者这时很难确定是发送信息还是接收信息。
请考虑如何能确定信息是不是重放的信息。
10. 如何保护IC卡的安全？
11. 请画出带有时间戳的基于秘密密钥的身份验证过程。
12. 简述认证机构的严格层次结构模型的性质。
13. 证书管理由哪3个阶段组成？每个阶段包括哪些具体内容？
14. 简述使用密钥的身份认证的分类方法。
15. 简述Kerberos身份认证的异域认证过程。
16. 简述X.509证书包含的内容。
17. 简述X.509的双向认证过程。
18. 试述数字证书的原理。
19. 查阅资料，简述有关PKI的标准及其相关产品。
20. PKI可以提供哪些安全服务？PKI体系中包含了哪些与信任有关的概念？
21. 叙述基于X.509的数字证书在PKI中的作用。
22. 查阅资料，简述有关PKI的标准及其相关产品。
23. PKI可以提供哪些安全服务？PKI体系中包含了哪些与信任有关的概念？
24. 在信息系统内主体通常指什么？客体通常指什么？
25. 查找资料，分别给出几个自主访问控制、强制访问控制和基于角色的访问控制的实例。
26. 比较自主访问控制、强制访问控制和基于角色的访问控制。
27. 查找资料，说明还有哪些新的访问控制策略。

# 第 4 章　网络安全防护

现代信息系统都是在网络环境下运行的。本章主要介绍有关网络环境的安全技术。

## 4.1　网络防火墙

在建筑群中，防火墙（firewall，见图 4.1）用来防止火灾蔓延。在计算机网络中，防火墙是设置在可信任的内部网络和不可信任的外界之间的一道屏障，阻滞不希望或者未授权的通信进出内部网络，通过强化边界控制来保障内部的安全，同时不妨碍内部对外部的访问，是目前实现网络安全的最有效的措施之一。



图 4.1　有防火墙的民居

### 4.1.1　网络防火墙概述

#### 1. 防火墙的作用

1）强化网络安全策略

防火墙是位于所保护网络边界上的关口，可以过滤数据包，可以选择符合规则的服务，对于来往的访问进行双向检查：能将可疑访问拒之门外，也能防止未经允许的访问到外部网络。当然，这就要求无论是从内部到外部的、还是从外部到内部的访问，都必须经过防火墙，并且只有被授权的通信才能通过防火墙，也可以防止内部信息外泄。

2）防止故障蔓延

由于防火墙具有双向检查功能，也能够将网络中一个网块（也称为网段）与另一个网块隔开，从而限制了局部重点或敏感网络安全问题对全局网络造成的影响，防止攻击性故障蔓延。

3）对网络访问进行监控审计和报警

防火墙位于网络的边界上，能有效地监控内部网和外部网之间的一切活动。当所有的访问都经过防火墙，防火墙就能记录下这些访问并做出日志。根据这些记录，管理人员就可以知晓网络的运行状况，知道网络是否受到了攻击，是什么样的攻击。当发生可疑动作时，防火墙能进行适当的报警，并提供网络是否受到监测和攻击的详细信息。

防火墙还可以具有分析功能，通过对有关记录的统计分析，知道网络有哪些威胁，有哪些安全需求，也能清楚防火墙是否能够抵挡攻击者的探测和攻击，并且清楚防火墙的控制是否充足。

4）提供流量控制（带宽管理）和计费

流量统计建立在流量控制基础之上。通过对基于 IP、服务、时间、协议等的流量进行

统计，可以实现与管理界面挂接，并便于流量计费。

流量控制分为基于 IP 地址的控制和基于用户的控制。基于 IP 地址的控制是对通过防火墙各个网络接口的流量进行控制；基于用户的控制是通过用户登录来控制每个用户的流量，防止某些应用或用户占用过多的资源，保证重要用户和重要接口的连接。

5）实现 MAC 与 IP 地址的绑定

MAC 与 IP 地址绑定起来，主要用于防止受控（不允许访问外网）的内部用户通过更换 IP 地址访问外网。这其实是一个可有可无的功能。不过因为它实现起来太简单了，内部只需要两个命令就可以实现，所以绝大多数防火墙都提供了该功能。

**2. 网络防火墙的局限**

1）防火墙有可能是可以绕过的

防火墙可以确定哪些内部服务允许外部访问，哪些外部用户可以访问所允许的内部服务，哪些外部服务可以由内部用户访问。为了发挥防火墙的作用，出入的信息必须经过防火墙，被授权的信息才能通过。因而，防火墙应当是不可渗透或绕过的，但若防火墙一旦被攻击者击穿或绕过，防火墙将失去作用。

实际上，系统往往会有缺陷，也往往会由于后门攻击而留下一些漏洞。如图 4.2 所示，如果内部网络中有一个未加限制的拨出，内部网络用户就可以（用向 ISP 购买等方式）通过串行链路网际协议（Serial Line Internet Protocol，SLIP）或点到点协议（Pointer-to-Pointer Protocol，PPP）与 ISP 直接连接，从而绕过防火墙。

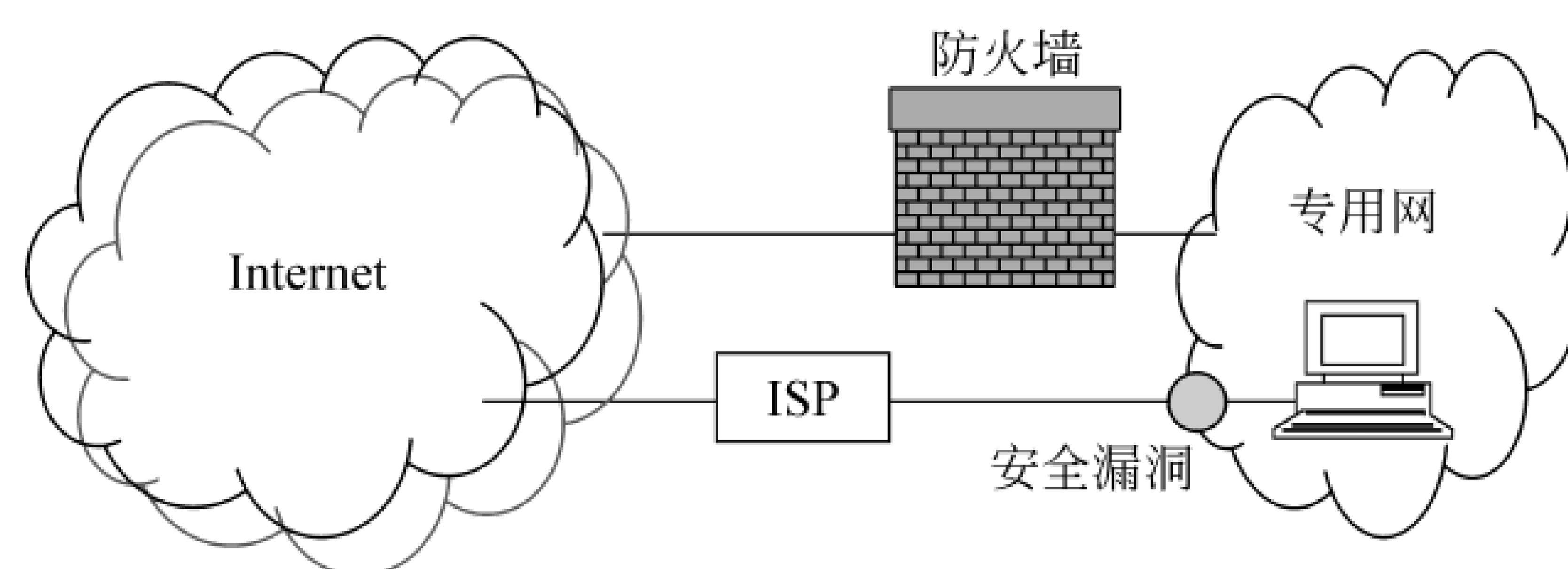



图 4.2　防火墙的漏洞

由于防火墙依赖于口令，所以防火墙不能防范黑客对口令的攻击。几年前，两个在校学生编了一个简单的程序，通过对波音公司的口令字的排列组合试出了开启内部网的钥匙，从网中搞到了一张授权的波音公司的口令表，将口令一一出卖。所以美国马里兰州的一家计算机安全咨询机构负责人诺尔•马切特说：“防火墙不过是一道较矮的篱笆墙。”黑客像耗子一样，能从这道篱笆墙上的窟窿中出入。这些窟窿常常是人们无意中留下来的，甚至包括一些对安全性有清醒认识的公司。例如，由于 Web 服务器通常处于防火墙体系之外，而有些公司随意扩展浏览器的功能，使之含有 Applet 编写工具。黑客们便可以利用这些工具钻空子，接管 Web 服务器，接着便可以从 Web 服务器出发溜过防火墙，大摇大摆地“回到”内部网中，好像他们是内部用户，刚刚出来办完事又返回去一样。

2）防火墙不能防止内部出卖性攻击或内部误操作

显然，当内部人员将敏感数据或文件复制到 U 盘等移动存储设备上提供给外部攻击者时，防火墙是无能为力的。此外，防火墙也不能防范黑客，黑客有可能伪装成管理人员或新职工，以骗取没有防范心理的用户的口令，或借用他们的临时访问权限实施攻击。

3）防火墙不能防止对开放端口（服务）的攻击

防火墙要保证服务，必须开放相应的端口。防火墙要准许 HTTP 服务，就必须开放 80 端口；要提供 MAIL 服务，就必须开放 25 端口等。防火墙不能防止对开放的端口进行攻击，即不能防止利用开放服务流入的数据攻击、利用开放服务的数据隐蔽隧道的攻击和对于开放服务软件缺陷的攻击。

4）防火墙不能防止数据驱动式的攻击

有些数据表面上看起来无害，可是当它们被邮寄或复制到内部网的主机中后，就可能会发起攻击，或为其他入侵准备好条件。这种攻击就称为数据驱动式攻击。防火墙无法防御这类攻击。

5）防火墙可以阻断攻击，但不能消灭攻击源

防火墙是一种被动防卫机制，不是主动安全机制。Internet 上的各种攻击源源不断。设置得当的防火墙可以阻挡它们，但是无法清除这些攻击源。

6）防火墙有可能自身遭到攻击

防火墙不能干涉还没有到达防火墙的包，如果这个包是攻击防火墙的，只有已经发生了攻击，防火墙才可以对抗。并且防火墙也是一个系统，也有自己的缺陷，也会受到攻击。这时许多防御措施就会失灵。

**3. 防火墙的种类**

从不同的角度，可以对防火墙进行不同的分类。下面介绍几种重要的防火墙分类。

1）按照采用的技术分类

按照采用的技术可将防火墙分为以下 4 类。

（1）地址转换防火墙。内部地址与外部地址不一致，可以防止外部直接根据地址进行攻击。

（2）数据包过滤防火墙。这种防火墙主要工作在 IP 层和 TCP 层，按照数据包的内容进行检查，按照默认允许或默认禁止两种原则进行过滤。

（3）代理防火墙。这种防火墙能够将所有跨越防火墙的网络通信链路分为两段，使得网络内部的用户不直接与外部的服务器通信，可以避免数据驱动式攻击。

（4）状态检测防火墙。是第三代防火墙技术，能对网络通信的各层实行检测。

2）按照实现形态分类

按照实现形态可将防火墙分为以下 3 类。

（1）软件防火墙。软件防火墙单独使用软件系统来完成防火墙功能，将软件部署在系统主机中的公共操作系统上。它要占用系统资源，在一定程度上影响系统性能。其一般用于单机系统或是极少数的个人计算机，很少用于计算机网络中，所以也称为个人防火墙。

（2）硬件防火墙。通常称为网络防火墙，其基本原理是把软件防火墙嵌入在硬件中，或者说是在一台服务器上装了软件防火墙。高端的硬件防火墙不是将防火墙软件装在硬盘中，而是固化在 BIOS 中，这样提高发包效率，并且很难被破坏（因为 BIOS 是 ROM 存储器，是只读型的）。

（3）芯片级防火墙。基于专门的硬件平台，没有操作系统。专有的 ASIC 芯片促使它们比其他种类的防火墙速度更快，处理能力更强，性能更高。做这类防火墙最出名的厂商有 NetScreen、FortiNet、Cisco 等。这类防火墙由于使用专用 OS（操作系统），因此防火墙本身的漏洞比较少，不过价格比较高昂。

3）从防火墙结构上分类

从结构上可将防火墙分为以下 3 类。

（1）单一主机防火墙。

（2）路由器集成式防火墙。

（3）分布式防火墙。

4）按防火墙的应用部署位置分类

按应用部署位置可将防火墙分为以下 3 类。

（1）边界防火墙。

（2）个人防火墙。

（3）混合防火墙。

5）按防火墙性能分类

按照性能可将防火墙分为以下两类。

（1）百兆级防火墙。

（2）千兆级防火墙。

### 4.1.2 防火墙技术之一——网络地址转换

网络地址转换（Network Address Translation，NAT）就是使用两套 IP 地址——内部 IP 地址（也称为私有 IP 地址）和外部 IP 地址（也称为公共 IP 地址）。私有 IP 地址是指内部网络或主机的 IP 地址，公有 IP 地址是指在 Internet 上全球唯一的 IP 地址。当受保护的内部网连接到 Internet 并且有用户要访问 Internet 时，它首先使用自己网络的内部 IP 地址，到了 NAT 后，NAT 就会从公共 IP 地址集中选一个未分配的地址分配给该用户，该用户即可使用这个合法的 IP 地址进行通信。同时，对于内部的某些服务器，如 Web 服务器，网络地址转换器允许为其分配一个固定的合法地址。外部网络的用户就可通过 NAT 来访问内部的服务器。

NAT 可以缓解 IP 地址较少与主机数量过多之间的矛盾，可以用少量的与数目较多的

内部主机之间相互映射。由于它对外隐藏内部主机的私有 IP 地址，也提高了内部网络的安全性。

虽然 NAT 可以借助于某些代理服务器来实现，但考虑到运算成本和网络性能，很多时候都是在路由器或防火墙上实现的。

### 1. 私有 IP 地址空间

RFC 1918 为私有网络预留了 3 个 IP 地址块，如下。

A 类：10.0.0.0～10.255.255.255。

B 类：172.16.0.0～172.31.255.255。

C 类：192.168.0.0～192.168.255.255。

上述 3 个范围内的地址不会在 Internet 上被分配，可以不必向因特网服务提供商（Internet Service Provider，ISP）或注册中心申请即可在公司或企业内部自由使用。

### 2. 基本 NAT 与 NAPT

NAT 技术于 20 世纪 90 年代提出。原来仅仅进行 IP 地址映射，后来把映射范围扩大到了端口，于是形成两种 NAT：基本 NAT 和 NAPT。

1）基本 NAT

基本 NAT 常被简称为 NAT，其特点是仅仅进行内部 IP 与公共 IP 之间的映射，不进行端口的映射。其特点是内部 IP 与公共 IP 具有一对一的映射关系。

基本 NAT 的实现有两种方法：静态 NAT（static NAT）和动态 NAT（dynamic NAT）。

静态 NAT 指私有 IP 地址与公共 IP 地址具有固定的一对一的映射关系。

**例 4.1** 假设内部局域网使用的 IP 地址段为 192.168.0.1～192.168.0.254，路由器局域网端（即默认网关）的 IP 地址为 192.168.0.1，子网掩码为 255.255.255.0。网络分配的公共 IP 地址范围为 61.159.62.128～61.159.62.135，路由器在广域网中的 IP 地址为 61.159.62.129，子网掩码为 255.255.255.248，可用于转换的 IP 地址范围为 61.159.62.130～61.159.62.134。于是可以得到表 4.1 所示的静态 NAT 表。

**表 4.1 例 4.1 的 NAT 表**

| 内部本地 IP | 内部全局 IP | 外部 IP |
|---|---|---|
| 192.168.0.2 | 200.168.12.9 | 61.159.62.130 |
| 192.168.0.3 | 200.168.12.3 | 61.159.62.131 |
| 192.168.0.4 | 200.168.12.5 | 61.159.62.132 |
| 192.168.0.5 | 200.168.12.1 | 61.159.62.133 |
| 192.168.0.6 | 200.168.12.6 | 61.159.62.134 |

实际上，NAT 就是内部本地地址与内部全局地址之间的转换。那么，只有内部本地 IP 地址，没有内部全局 IP 地址是否可以呢？如果这样，所提供的 NAT 表只能用于内部主机对于外部的访问，外部主机无法访问内部主机。为了能让外部主机访问内部主机，就要向外

公布内部主机的 IP 地址。但这样就失去设置内部 IP 地址的意义了。为每一台内部主机增加一个全局 IP 地址的目的是让外部主机只能知道这个外部地址，无法知道真实的内部地址。要访问内部主机必须经过防火墙检查，从而为内部主机提供了一层保护。当然，这时，必须申请与主机数同样多的内部全局 IP 地址。

动态 NAT 指私有 IP 地址与公共 IP 地址具有动态（临时）的一对一映射关系。基本方法是将可用的全局地址集定义成 NAT 池（NAT pool）。对于要与外界进行通信的内部节点，如果还没有建立转换映射，边缘路由器或者防火墙将会动态地从 NAT 池中选择全局地址对内部地址进行转换。每个转换条目在连接建立时动态建立，而在连接终止时会被回收。这样，网络的灵活性大大增强了，所需要的全局地址进一步减少，所以动态 NAT 适合于内部主机数大于全局 IP 地址数的情形。

动态转换提供了很大的灵活性，但增加了网络管理的复杂性。特别是当 NAT 池中的全局地址被全部占用以后，以后的地址转换的申请会被拒绝，所以内部主机同时访问外部主机的数目取决于申请到的内部全局 IP 地址的数目。这样会造成网络连通性的问题，一般只用于拨号连接或频繁的远程连接中。

2）NAPT

网络地址端口转换（Network Address Port Translation，NAPT）不仅进行 IP 地址映射，还进行 TCP 或 UDP 端口映射，即进行的是<内部地址+内部端口>与<外部地址+外部端口>之间的映射。NAPT 主要采用地址端口转换的方法进行转换，即将内部地址映射到外部网络的 IP 地址的不同端口上，从而可以实现多对一的映射。表 4.2 为一个 NAPT 表的示例，其 3 台内部主机共用了一个全局 IP 地址。

**表 4.2　NAPT 表的示例**

| 协　　议 | 内部本地地址端口 | 内部全局地址端口 | 外部全局地址端口 |
|---|---|---|---|
| TCP | 10.1.1.3：1723 | 202.168.2.2：1492 | 212.21.7.3：23 |
| TCP | 10.1.1.2：1723 | 202.168.2.2：1723 | 212.21.7.3：23 |
| TCP | 10.1.1.1：1034 | 202.168.2.2：2034 | 212.21.7.3：23 |

### 3．NAT 的优缺点

NAT 使内部网络的计算机不可能直接访问外部网络：通过包过滤分析，若传入的包没有专门指定配置到 NAT，就将之丢弃。同时使所有内部的 IP 地址对外部是隐蔽的。因此，网络之外没有谁可以通过指定 IP 地址的方式直接对网络内的任何一台特定的计算机发起攻击。NAT 还可以使多个内部主机共享数量有限的 IP 地址。还可以启用基本的包过滤安全机制，NAT 虽然可以保障内部网络的安全，但也有一些局限。例如，内部用户可以利用某些木马程序通过 NAT 做外部连接。

### 4.1.3 防火墙技术之二——代理服务

#### 1. 代理服务器

代理服务器（proxy server）是用户计算机与 Internet 之间的中间代理机制，它采用客户/服务器工作模式。代理服务器位于客户与 Internet 上的服务器之间。请求由客户端向服务器发起，但是这个请求要首先被送到代理服务器；代理服务器分析请求，确定其是合法的以后，首先查看自己的缓存中有无要请求的数据，有就直接传送给客户端，否则再以代理服务器作为客户端向远程的服务器发出请求；远程服务器的响应也要由代理服务器转交给客户端，同时代理服务器还将响应数据在自己的缓存中保留一份副本，以备客户端下次请求时使用。图 4.3 为代理服务的结构及其数据控制和传输过程示意图。

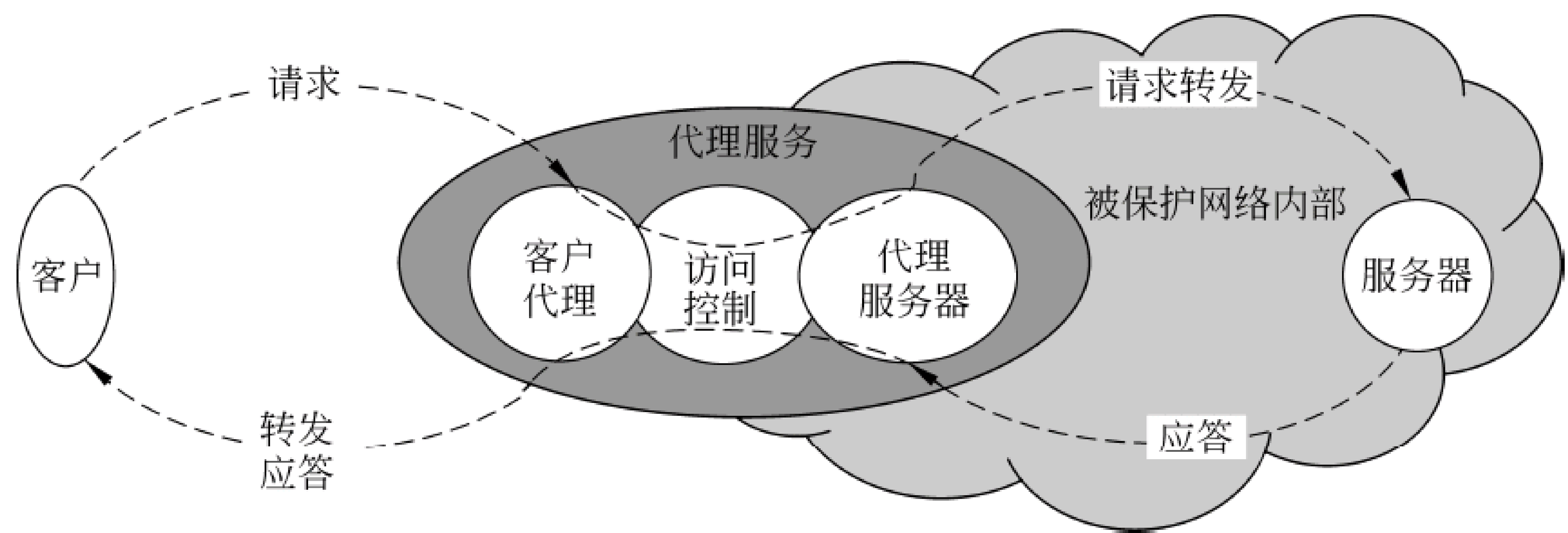



图 4.3　代理服务的结构及其数据控制和传输过程

应用于网络安全的代理技术，也是要建立一个数据包的中转机制，并在数据的中转过程中加入一些安全机制。

代理技术可以在不同的网络层次上进行。主要的实现层次在应用层和传输层，分别称为应用级代理和电路级代理。它们的工作原理有所不同。

#### 2. 应用级代理

应用级代理像横在客户与服务器连通路径上的一个关口，所以也被称为应用层网关（application level gateway）。由于应用级代理还像横在客户与服务器连通路径上的一堵墙，所以也被称为应用级防火墙。如图 4.4 所示，应用级代理只有为特定的应用程序安装了代理程序代码，该服务才会被支持，并建立相应的连接。显然，这种方式可以拒绝任何没有明确配置的连接，从而提供了额外的安全性和控制性。但是，应用级代理没有通用的安全机制和安全规则描述，它们通用性差，对不同的应用具有很强的针对性和专用性。

图 4.5 为应用级代理的基本工作过程。

应用级代理具体提供如下一些功能。

1）阻断路由与 URL

代理服务是一种服务程序，它位于客户机与服务器之间，完全阻挡了两者间的数据交流。从客户机来看，代理服务器相当于一台真正的服务器；而从服务器来看，代理服务器又是一台真正的客户机。当客户机需要使用服务器上的数据时，首先将数据请求发给代理

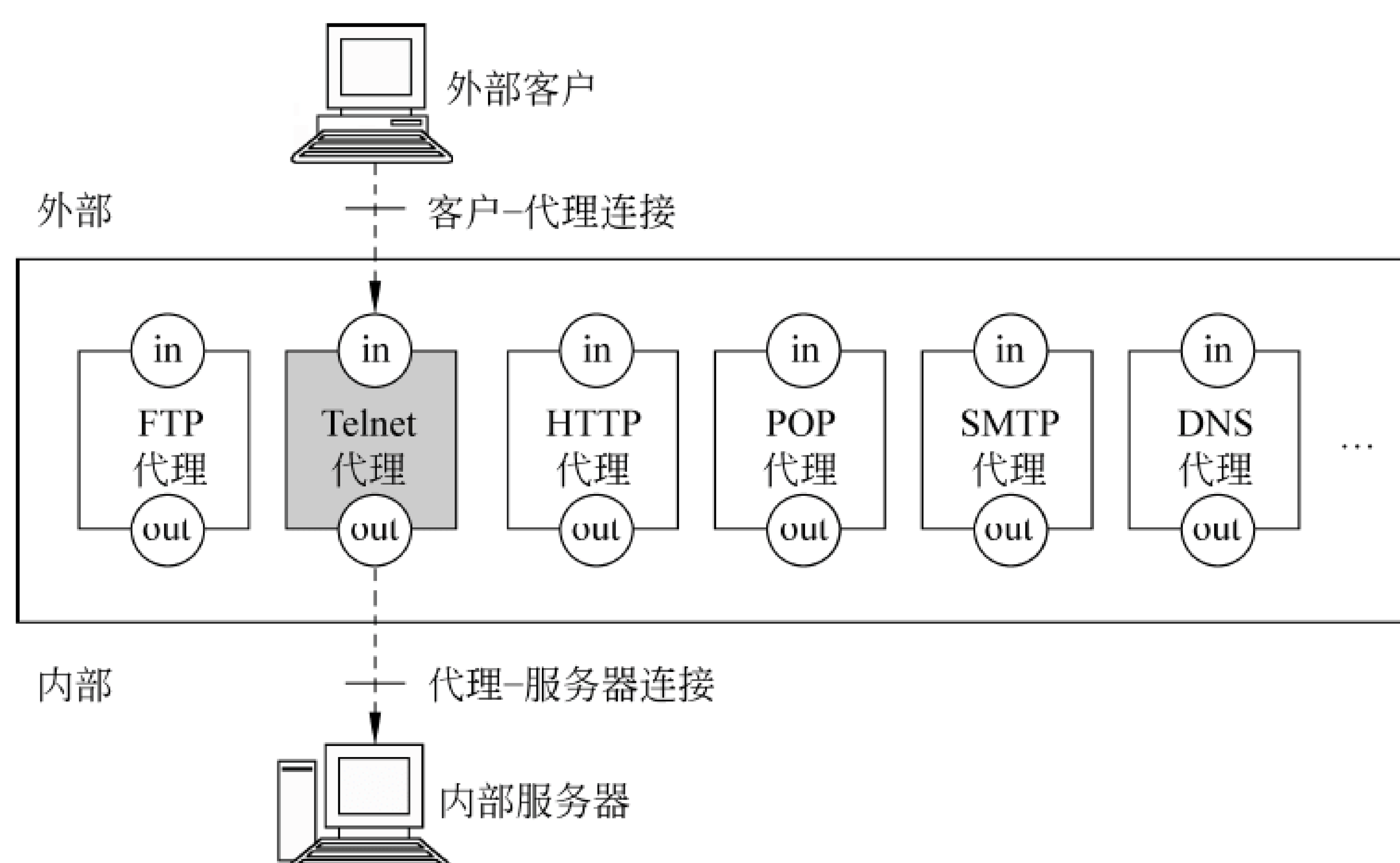



图 4.4　应用级代理工作原理

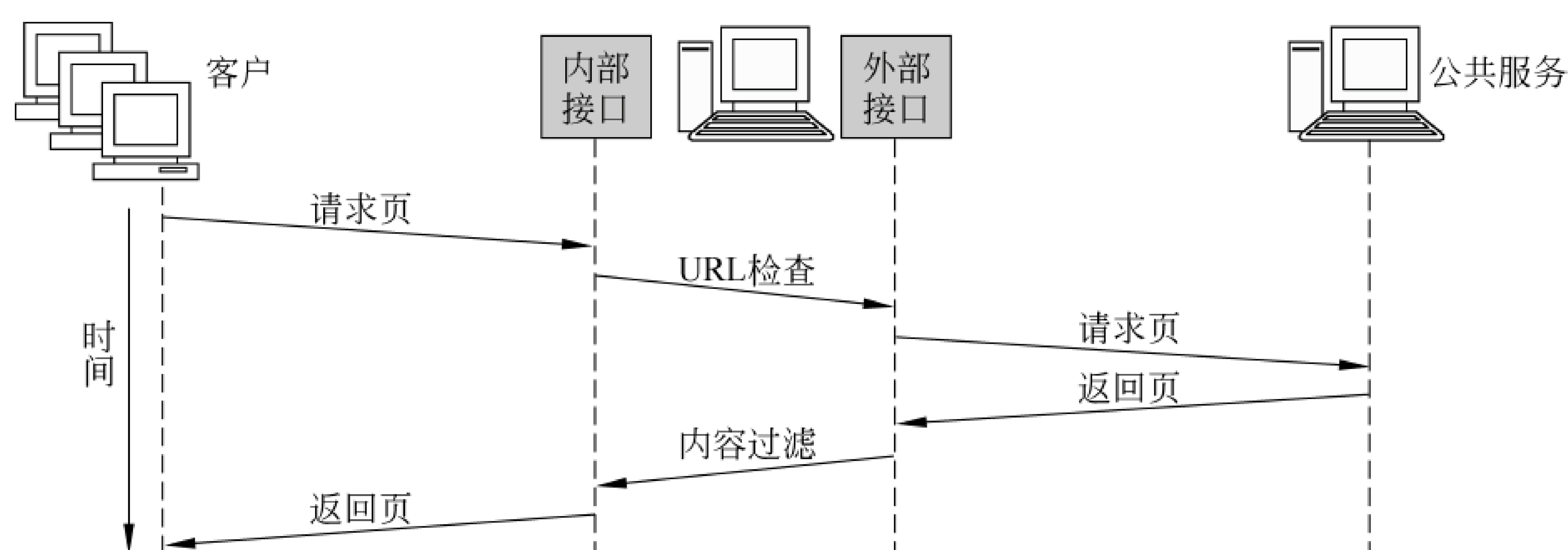



图 4.5　应用级代理的基本工作过程

服务器，代理服务器再根据这一请求向服务器索取数据，然后再由代理服务器将数据传输给客户机。由于外部系统与内部服务器之间没有直接的数据通道，外部的恶意侵害也就很难伤害到企业内部网络系统。

代理通过侦听网络内部客户的服务请求，然后把这些请求发向外部网络。在这一过程中，代理要重新产生服务级请求。例如，一个 Web 客户向外部发出一个请求时，这个请求会被代理服务器“拦截”，再由代理服务器向目标服务器发出一个请求。因此外部主机与内部主机之间并不存在直接连接，从而可以防止传输层因源路由、分段和不同的服务拒绝造成的攻击，确保没有建立代理服务的协议不会被发送到外部网络。

2）隐藏用户

应用级代理既可以隐藏内部 IP 地址，也可以给单个用户授权，即使攻击者盗用了一个合法的 IP 地址，也通不过严格的身份认证。因此应用级代理比数据包过滤具有更高的安全性。但是这种认证使得应用网关不透明，用户每次连接都要受到认证，这给用户带来许多不便。这种代理技术需要为每个应用写专门的程序。

代理保证所有内容都经过单一的一个点，该点成为网络数据的一个检查点。在应用级代理提供授权检查及代理服务。大多数代理软件可以具有对过往的数据包进行分析监控、

注册登记、过滤、记录和报告等功能，当外部某台主机试图访问受保护网络时，必须先在代理上经过身份认证。通过身份认证后，再运行一个专门为该网络设计的程序，把外部主机与内部主机连接。在这个过程中，可以限制用户访问的主机、访问时间及访问的方式进行记录、监控。同样，受保护网络内部用户访问外部网时也需要先登录到代理上，通过验证后，才可访问。当发现被攻击迹象时会向网络管理员发出警报，并能保留攻击痕迹。代理服务器的缺点是必须针对客户机可能产生的所有应用类型逐一进行设置，大大增加了系统管理的复杂性。此外，假如由于黑客攻击等原因使代理不工作时，对应的服务请求也就被切断了。这也是单点访问的不足之处。

3）提高用户访问效率

代理服务器起内容中转的作用，能把服务器向用户提供的内容先保存起来，再提供给用户，下次用户有同样请求时，就会只从服务器传输改变的部分，从而提高了用户访问效率。

### 3. 电路级代理

电路级代理也称为电路层网关（circuit level gateway）。如图 4.6 所示，在 OSI 模型中电路层网关工作在会话层，它维护一张合法的会话连接表，进行会话层的过滤。在 TCP/IP 协议栈中，电路层网关在 TCP 三次握手过程中检查双方的 SYN、ASK 和序列号是否合乎逻辑，依此判断请求的会话是否合法。一旦认为会话合法，就为双方建立连接。之后就只作为数据包的中转站，进行简单的字节复制式的数据包转接，不再进行任何审查、过滤和管理。因此，受保护网与外部网的信息交换是透明的。

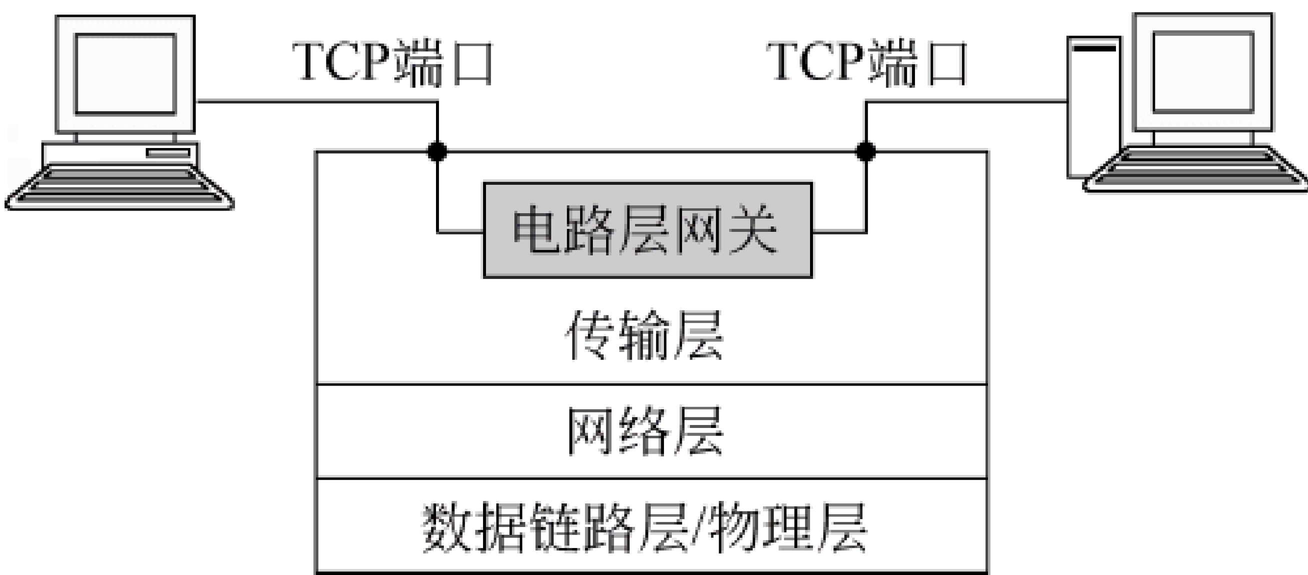



图 4.6　电路层网关工作原理

处于安全网络内的客户端可以事先通知电路层网关有哪些包要到来，这也就形成了一个隐患，即内部用户为外部主机设置了一条特殊通道。为此，电路层网关要与过滤型防火墙一起使用，并要作好日志。

### 4. SOCKS 协议

1）SOCKS 协议的组成

SOCKS 协议（套接字协议）是一个电路层网关协议，它主要由两部分组成。

（1）SOCKS 客户程序：经过修改的 Internet 客户程序。改造的目的是使运行客户程序的主机从与 Internet 通信改为与运行 SOCKS 代理的主机通信。

（2）SOCKS 服务程序：既可以与 Internet 通信又可以与内部网络通信的程序。

2）SOCKS 代理的工作过程

（1）当一个经过 SOCKS 化的客户程序要连接到 Internet 时，SOCKS 就会截获这个连接，将之连接到运行 SOCKS 服务器的主机上。

（2）连接建立后，SOCKS 客户程序发送如下信息。

① 版本号。

② 连接请求命令。

③ 客户端端口号。

④ 发起连接的用户名。

（3）经过确认后，SOCKS 服务器才与外部的服务器建立连接。

### 4.1.4　防火墙技术之三——包过滤

#### 1. 数据包及其结构

在网络中传输的数据是从应用程序那里递交来的。应用程序递交给网络要传输的数据后，网络就要逐层向下，转交给下面的一层去实施，每交到下一层，就要按照本层的协议要求进行一次打包，形成不同协议层中的数据包（packet），直到物理网络。图 4.7 表明在 TCP/IP 网络中数据包的封装与解包过程。图中的虚箭线为发送端的数据封装过程，实箭线表示接收端的数据解包过程。

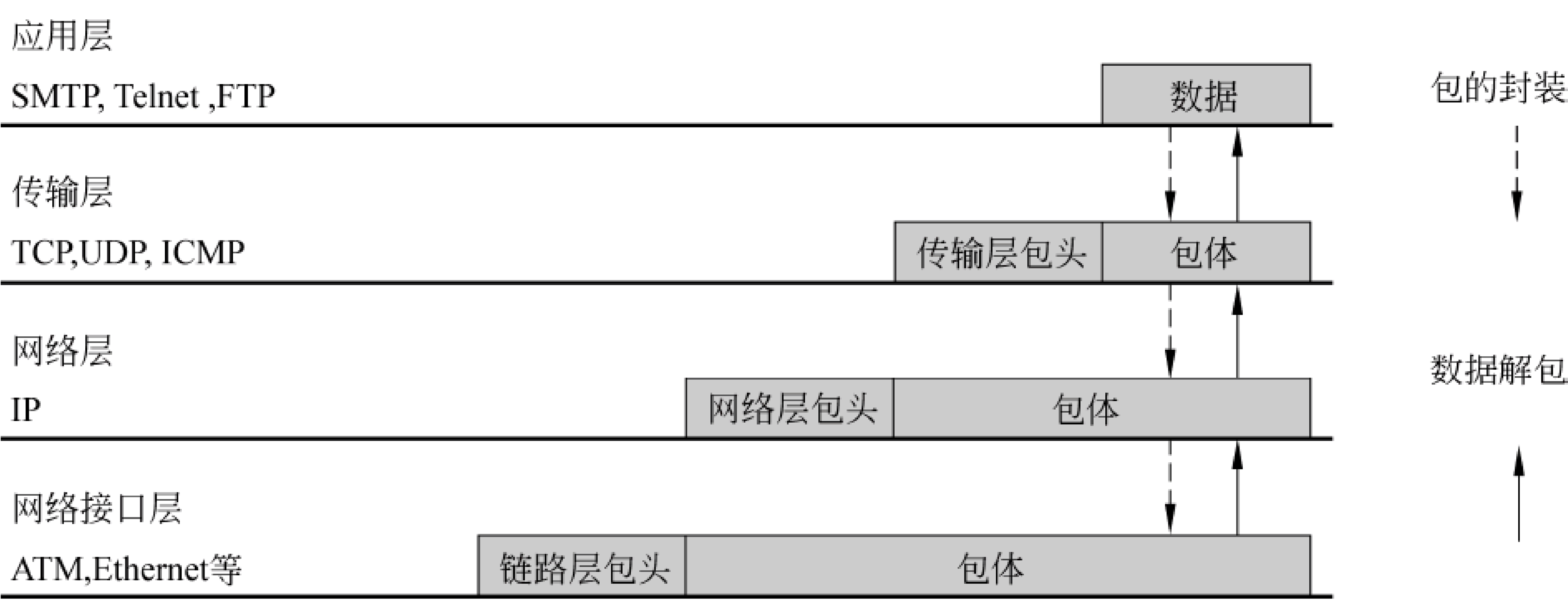



图 4.7　TCP/IP 网络中数据包的封装与解包

应当注意，包过滤是根据数据包的特征进行的，其中主要根据数据包头的一些字段的特征进行。由于不同的协议所规定的包头格式不同，因此在制定过滤规则前，应当充分了解数据包的格式。图 4.8 中列出了其他一些常用的数据包的格式，供本书后面的讨论中使用。

下面介绍这些数据包中可以体现数据包特征的有关字段。

（1）源地址（Source Address）和目的地址（Destination Address）：它们各表明数据包的源 IP 地址和目的 IP 地址。根据地址，还可以判断出数据流的方向：是由外部网络流入内部网络——往内（流入），还是由内部网络流入外部网络——往外（流出）。

（2）标识符：是发送方分配的一个独一无二的编号，用于标识同一数据报中的各分组，以便组装。

| 0–3 | 4–7 | 8–15 | 16–18 | 19–31 |
|---|---|---|---|---|
| 版本 | 头标长 | 服务类型 | 总长 | |
| 标识 | | | 标志 | 片偏移 |
| 生存时间 | | 协议 | 报头校验和 | |
| 源IP地址 | | | | |
| 目的IP地址 | | | | |
| IP分组选项 | | | | 填充 |
| 数据 | | | | |

(a) IP分组格式

| | |
|---|---|
| UDP源端口 | UDP目的端口 |
| UDP数据报长度 | UDP校验和 |
| UDP数据区 | |
| … | |

(b) UDP数据报格式

| 0 … 7 | 8 … 15 | |
|---|---|---|
| 类型码 | 代码 | 校验和 |
| 首部其余部分 | | |
| 数据部分 | | |

(c) ICMP分组的格式

图 4.8　一些数据包的格式

（3）源端口（Source Port）和目的端口（Destination Port）：在 TCP 和 UDP 数据包中，源端口和目的端口分别表示本地通信端口和异地通信端口。端口号是按照协议类型分配的，所以端口号也表明了所传输的数据包服务的协议类型。

（4）协议（Protocol）：在 IP 数据包中，“协议”字段用于标识接收的 IP 分组中的数据的高层（传输层）协议。高层协议号由 TCP/IP 协议中央权威机构 NIC（Network Information Center）分配，例如，1 为控制报文协议 ICMP，6 为传输控制协议 TCP，8 为外部网关协议 EGP，17 为用户数据报协议 UDP，29 为传输层协议第 4 类 ISO-TP4。

（5）服务类型（Type of Service，ToS）：在 IP 数据包中，ToS 描述 IP 分组所希望获得的服务质量，占 8 位，包括如下几项。

① 低延迟、高吞吐量、高可靠性，各占 1b。

② 优先级，共 8 级，占 3b。

③ 未用 2b。

表 4.3 列出了 RFC 1349[Almquist 1992]对于不同应用建议的 ToS 值。

**表 4.3　RFC 1349[Almquist 1992]对于不同应用建议的 ToS 值**

| 应 用 程 序 | | 最小时延 | 最大吞吐量 | 最高可靠性 | 最小费用 | 十六进制值 |
|---|---|---|---|---|---|---|
| Telnet/Rlogin | | 1 | 0 | 0 | 0 | 0x10 |
| FTP | 控制 | 1 | 0 | 0 | 0 | 0x10 |
| | 数据 | 0 | 1 | 0 | 0 | 0x08 |
| | 任意块数据 | 0 | 1 | 0 | 0 | 0x08 |
| TFTP | | 1 | 0 | 0 | 0 | 0x10 |
| SMTP | 命令 | 1 | 0 | 0 | 0 | 0x10 |
| | 数据 | 0 | 1 | 0 | 0 | 0x08 |

续表

| 应 用 程 序 | | 最小时延 | 最大吞吐量 | 最高可靠性 | 最小费用 | 十六进制值 |
|---|---|---|---|---|---|---|
| DNS | UDP 查询 | 1 | 0 | 0 | 0 | 0x10 |
| | TCP 查询 | 0 | 0 | 0 | 0 | 0x00 |
| | 区域传输 | 0 | 1 | 0 | 0 | 0x08 |
| ICMP | 差错 | 0 | 0 | 0 | 0 | 0x00 |
| | 查询 | 0 | 0 | 0 | 0 | 0x00 |
| | 任何 IGP | 0 | 0 | 1 | 0 | 0x04 |
| SNMP | | 0 | 0 | 1 | 0 | 0x04 |
| BOOTP | | 0 | 0 | 0 | 0 | 0x00 |
| NNTP | | 0 | 0 | 0 | 1 | 0x02 |

（6）数据包内容：前面的 5 个字段都来自数据包头中，而数据内容则是来自数据包体中。例如，数据内容中一些关键词可以代表数据内容的某一方面的特征。对数据包内容的抽取将会形成依据内容的包过滤（packet filtering）规则。这是目前包过滤技术研究的一个重要方面。

### 2. 包过滤安全策略的制定

早期的包过滤是在路由器上进行的。通过对路由表的配置来决定数据包是否符合过滤规则。数据包的过滤规则（即访问控制列表，ACL）由一些规则逻辑描述：一条过滤规则规定了允许数据包流进或流出内部网络的一个条件。

在制定了数据包过滤规则后，对于每一个数据包，路由器会从第一条规则开始逐条进行检查，最后决定该数据包是否符合过滤逻辑。

包过滤的核心技术是安全策略和过滤原则的设计。其大致步骤如下。

（1）进行安全需求分析，确定安全策略。根据网络的具体情况，确定需要保护什么，需要提供什么服务，进一步明确所允许和禁止的任务。

（2）将安全策略转化为一个数据包过滤规则表。过滤规则的设计主要依赖于数据包中的信息：源地址、目标地址、TCP/UDP 源端口号、TCP/UDP 目的端口号、标志位、协议以及内容等。

制定包过滤规则的一条基本原则是“最小特权原则”。从安全的角度，默认拒绝应该更可靠。此外，包过滤还有禁入和禁出的区别。前者不允许指定的数据包由外部网络流入内部网络，后者不允许指定的数据包由内部网络流入外部网络。

此外，在制定过滤规则时，除了明确禁止和允许的规则外，还应考虑对明确的规则没有考虑到的其他情况的过滤规则。这些针对“其他”情况的规则称为默认规则。默认规则可以采用如下原则之一。

① 默认接受（转发）：凡未被禁止的，就是允许的，即除明确指定禁止的数据包，其他都是允许通过（转发）的。这也称为“黑名单”策略。

② 默认拒绝（丢弃）：凡未被允许的，就是禁止的，即除明确指定通过的数据包，其他都是被禁止（丢弃）的。这也称为“白名单”策略。

（3）用过滤规则语法描述过滤逻辑。

（4）按照过滤逻辑对路由器进行设置。

### 3. 路由器进行包过滤的过程

（1）包过滤规则必须被包过滤设备端口存储起来。应用于包的规则顺序与包过滤器规则的存储顺序必须相同。

（2）当包到达端口时，对包头进行语法分析。大多数包过滤设备只检查 IP、TCP 或 UDP 报头中的字段。

（3）按照以下规则进行过滤操作。

① 若一条规则阻止包传输或接收，则此包便不被允许。

② 若一条规则允许包传输或接收，则此包便可以被继续处理。

③ 若没有一条规则匹配，就执行默认操作。

显然，过滤效果与规则的排列顺序有关。

### 4. 基于地址的数据包过滤策略

按照地址进行过滤是最简单的过滤方式，它的过滤规则只对数据包的源地址、目的地址和地址偏移量进行判断，这在路由器上是非常容易配置的。对于信誉不好或内容不宜并且地址确定的主机，用这种策略通过简单配置就可以将之拒之门外。但是，对于攻击尤其是地址欺骗攻击的防御，过滤规则的配置就要复杂多了。下面分几种情形分别考虑。

1）针对 IP 源地址欺骗攻击的配置

对于攻击者伪装内部用户的 IP 地址攻击，可以按照下面的原则配置过滤规则：如果发现具有内部地址的数据包到达路由器的外部接口，就将其丢弃。

显然，这种规则对于外部主机冒充另外一台主机的攻击则无能为力。

2）针对源路由攻击的配置

攻击者有时为了躲过网络的安全设施，要为数据包指定一个路由，这条路由可以使数据包以不期望路径到达目标。对付这种攻击的过滤规则是丢弃所有含有源路由的数据包。

3）针对小分片攻击的配置

当一个 IP 包太长时，就要对其进行分片传输。分组后，传输层的首部只出现在 IP 层的第 1 片中。攻击者利用 IP 分片的这一特点，往往会建立极小的分片，希望过滤路由器只检查第 1 片，而忽略后面的分组。

对付小分段攻击的策略是丢弃 FO 为 1 的 TCP、UDP 数据包。

**例 4.2**　某公司有一个 B 类网（123.45）。该网的子网（123.45.6.0/24）有一个合作网络（135.79）。管理员希望：

（1）禁止一切来自 Internet 的对内网的访问。

（2）允许来自合作网络的所有子网（135.79.0.0/16）访问内网（123.45.6.0/24）。

（3）禁止对合作网络的子网（135.79.99.0/24）的访问权（对全网开放的特定子网除外）。

为简单起见，只考虑从合作网络流向公司的数据包，对称地处理逆向数据包只需互换规则行中源地址和目的地址即可。表 4.4 为其网络 ACL。其中，规则 C 是默认规则。

**表 4.4　某公司网络的 ACL**

| 规　　则 | 源地址 | 目 的 地 址 | 过滤操作 |
|---|---|---|---|
| A | 135.79.0.0/16 | 123.45.6.0/24 | 允许 |
| B | 135.79.99.0/24 | 123.45.0.0/16 | 拒绝 |
| C | 0.0.0.0/0 | 0.0.0.0/0 | 拒绝 |

表 4.5 是使用一些样本数据包对表 4.4 所示过滤规则的测试结果。需要注意的是，规则的执行顺序对于执行结果有很大的影响。

**表 4.5　使用样本数据包测试结果**

| 数据包序号 | 源地址 | 目的地址 | 目标行为操作 | ABC 行为操作 | BAC 行为操作 |
|---|---|---|---|---|---|
| 1 | 135.79.99.1 | 123.45.1.1 | 拒绝 | 拒绝（B） | 拒绝（B） |
| 2 | 135.79.99.1 | 123.45.6.1 | 允许 | 允许（A） | 拒绝（B） |
| 3 | 135.79.1.1 | 123.45.6.1 | 允许 | 允许（A） | 允许（A） |
| 4 | 135.79.1.1 | 123.45.1.1 | 拒绝 | 拒绝（C） | 拒绝（C） |

可见，按 ABC 的规则顺序，能够得到想要的操作结果；而按 BAC 的规则顺序则得不到预期的操作结果，原本允许的数据包 2 被拒绝了。仔细分析可以发现，表 4.3 中用来禁止合作网的特定子网的访问规则 B 是不必要的。它正是在 BAC 规则集中造成数据包 2 被拒绝的原因。如果删除规则 B，得到表 4.6 所示的行为操作。

**表 4.6　删除规则 B 后的行为操作**

| 数据包 | 源地址 | 目的地址 | 目标行为操作 | AC 行为操作 |
|---|---|---|---|---|
| 1 | 135.79.99.1 | 123.45.1.1 | 拒绝 | 拒绝（C） |
| 2 | 135.79.99.1 | 123.45.6.1 | 允许 | 允许（A） |
| 3 | 135.79.1.1 | 123.45.6.1 | 允许 | 允许（A） |
| 4 | 135.79.1.1 | 123.45.1.1 | 拒绝 | 拒绝（C） |

这才是想要的结果。由此得出两点结论。

（1）正确地制定过滤规则是困难的。

（2）过滤规则的重新排序使得正确地指定规则变得越发困难。

### 5. 基于服务的数据包过滤策略

按服务进行过滤，就是根据 TCP/UDP 的端口号制定过滤策略。但是，由于源端口是可以伪装的，所以基于源端口的过滤是会有风险的。同时还需要确认内部服务确实是在相应的端口上。下面进行一些分析。

1）关于外部服务的端口号

如果过滤规则完全依赖于外部主机的端口号，例如，允许内部主机向外部服务器的邮

件发送服务，而且 TCP 的端口 25 就是常规邮件（STMP）端口时，这样的配置是安全的。但是，包过滤路由器是无法控制外部主机上的服务确实在常规的端口上，攻击者往往会通过伪造，利用端口 25 向内部主机发送其他应用程序（非常规邮件）的数据包，建立连接，进行非授权访问。这时，只能禁止 25 端口对于内部主机的访问。因为内部主机对这个外部端口不能信任。

2）关于内部主机的源端口号

从内部到外部的 TCP/UDP 连接中，内部主机的源端口一般采用大于 1024 的随机端口。为此，对端口号大于 1024 的所有返回到内部的数据包都要允许，不过，还需要辨认端口号大于 1024 的数据包中哪些是伪造的。

对于 TCP 数据包来说，可以通过 flag 位辨认哪些是来自外部的连接请求。但是 UDP 是无连接的，没有这样的 flag 位可使用，只能辨认端口号。所以允许 UDP 协议对外访问会带来风险，因为返回的数据包上的端口号有可能是攻击者伪造的。当请求端口和目的端口都固定时，这个问题才能解决。

**例 4.3** 表 4.7 与表 4.8 就是否考虑数据包的源端口进行对照。表 4.7 所示的规则表由于未考虑到数据包的源端口，出现了两端所有端口号大于 1024 的端口上的非预期的作用。而表 4.8 所示的规则表考虑到数据包的源端口，所有规则限定在 25 号端口上，故不可能出现两端端口号均在 1024 以上的端口上连接的交互。

**表 4.7 未考虑源端口时的包过滤规则**

| 规则 | 方向 | 类型 | 源地址 | 目的地址 | 目的端口 | 行为操作 |
|---|---|---|---|---|---|---|
| A | 入 | TCP | 外 | 内 | 25 | 允许 |
| B | 出 | TCP | 内 | 外 | ≥1024 | 允许 |
| C | 出 | TCP | 内 | 外 | 25 | 允许 |
| D | 入 | TCP | 外 | 内 | ≥1024 | 允许 |
| E | 出/入 | 任何 | 任何 | 任何 | 任何 | 禁止 |

**表 4.8 考虑了源端口时的包过滤规则**

| 规则 | 方向 | 类型 | 源地址 | 目的地址 | 源端口 | 目的端口 | 行为操作 |
|---|---|---|---|---|---|---|---|
| A | 入 | TCP | 外 | 内 | ≥1024 | 25 | 允许 |
| B | 出 | TCP | 内 | 外 | 25 | ≥1024 | 允许 |
| C | 出 | TCP | 内 | 外 | ≥1024 | 25 | 允许 |
| D | 入 | TCP | 外 | 内 | 25 | ≥1024 | 允许 |
| E | 出/入 | 任何 | 任何 | 任何 | 任何 | 任何 | 禁止 |

### 4.1.5 防火墙技术之四——状态检测

状态检测（stateful-inspection）防火墙又称动态包过滤防火墙，是第三代防火墙技术，它通过安装在网关的软件——检查引擎，在不影响网络正常运行的前提下，截获数据包并抽取有关数据对网络的各层进行实时检测，跟踪每一个有效连接的状态，并根据这些信息

决定对该连接是接受还是拒绝。这种技术提供了高度安全的解决方案，同时具有较好的适应性和扩展性。

### 1. 静态数据包过滤的问题

相对于状态检测防火墙而言，前面介绍的数据包过滤也被称为无状态数据包过滤。因为它仅单独分析每一个数据包，不考虑包内高层的信息以及不同包之间的逻辑关系，也不关心数据传输的状态。这就会为攻击者提供机会。

**例 4.4** 一个防火墙的过滤规则为将目标端口≤1203 的数据包丢弃。这样，一位攻击者可以连续地发一系列伪装的 TCP ACK 包，每个包的端口分别为 1200、1201、1202、1203、1204。尽管这一系列包都是违背 TCP 协议的，因为发起连接必须是 SYN 包，但防火墙并不检测 TCP 的标志位，仅检测端口号，阻挡了前 3 个数据包，放过了端口号为 1204 的包。这个包到达主机后，主机依据 TCP 协议发现了问题，会发一个 RST 包通知发送者终止本次连接。不过，这恰中了攻击者的之计，使攻击者通过防火墙对内部某主机进行了一次半连接扫描攻击。

### 2. 状态检测防火墙的状态表

状态检测防火墙也称为动态包过滤防火墙，其检测引擎监视和跟踪每一个有效连接的状态，动态地维护一个状态信息表，通过规则表与状态表的共同配合，对表中的各个连接状态因素加以识别。下面分别介绍为 TCP 包和 UDP 包建立状态表的方法。

1）TCP 包

众所周知，一个 TCP 传输是在三次握手之后进行的。因此，可以把 TCP 包分为两类：握手包和传输数据包。它们的区别在包中的 6 个标志位上。当第一个带有 SYN 标志的数据包经过防火墙时，防火墙会根据报文的五元组信息（源 IP 地址、源端口、目的 IP 地址、目的端口、连接状态）检查自己的规则集，如果规则允许该报文通过，则把该报文的五元组信息写入状态表并根据路由表转发该报文。然后防火墙会设置一个超时时间，并等待服务器端 SYN/ACK 标志位为一的报文过来，如果在超时时间内防火墙收到了来自服务器的 SYN/ACK 标志位置一的数据报文，则状态表建立完成；如果在超时时间内未收到来自服务器的 SYN/ACK 报文，则状态表建立失败。

表 4.9 为状态检测防火墙状态表的一个实例。

**表 4.9 状态检测防火墙状态表的一个实例**

| 源 IP 地址 | 源端口号 | 目的 IP 地址 | 目的端口号 | 连接状态 |
|---|---|---|---|---|
| 192.168.1.100 | 1030 | 210.9.88.29 | 80 | 已建立 |
| 192.168.1.102 | 1031 | 216.32.42.123 | 80 | 已建立 |
| 192.168.1.101 | 1033 | 173.66.32.122 | 25 | 已建立 |
| 192.168.1.102 | 1035 | 177.231.32.12 | 79 | 已建立 |
| 223.43.21.231 | 1990 | 192.168.1.6 | 80 | 已建立 |

状态检测防火墙将属于同一连接的所有包作为一个整体的数据流看待，对于外部传来

的 TCP 数据包，首先检查它是否握手包，如果是握手包，就看其是否内网主机期待的包，是则允许，否则拒绝。如果不是握手包，则检查状态表中所记录的连接状态，如果属于已经建立的连接，则允许其通行，否则将其清除。因此，与简单包过滤相比，状态检测防火墙可以为包过滤提供更准确的信息，具有更高的安全性，但也消耗较多的计算资源。状态检测防火墙是一种基于状态检测的包处理器。

2）UDP 包

UDP 包比较简单，不携带任何连接或序列信息，只包含源 IP 地址、目的 IP 地址、源端口号、目的端口号、校验和以及所携带的数据。有些状态检测的设备也可以针对 UDP、ICMP 协议的交互过程建立状态表。但是这种状态表是以虚连接（virtual connection）为基础的，即将所有通过防火墙的 UDP 分组均视为一个虚连接，当反向应答分组送达时，就认为一个虚拟连接已经建立。如果在指定的一段时间内响应数据包没有到达，连接超时，则该连接被阻塞。所以状态检测技术最适合提供对 UDP 协议的有效支持。

**3. 状态检测防火墙的优点**

1）更高的安全性

实际上，现在状态检测防火墙并非仅仅抽取和检测传输层的有关数据，它工作在网关，在网络体系中处于数据链路层和网络层之间，从这里截取数据包，可以抽取和监测各层的有关数据。一般说来，状态检测防火墙首先在低协议层上检查数据包是否满足企业的安全策略，对于满足的数据包，再从更高协议层上进行分析。并且，它取到数据包后，首先根据安全策略从数据包中提取有用信息，保存在内存中；然后将相关信息组合起来，通过逻辑或数学运算来决定对数据包进行什么样的操作：允许通过、拒绝通过、认证连接、加密数据等。这样安全性得到很大提高。

2）更高的效率

由于状态检测防火墙工作在协议栈的较低层，通过防火墙的所有数据包都在低层处理，而不需要协议栈的上层处理任何数据包，这样减少了高层协议头的开销，执行效率提高很多。其次，状态检测防火墙不要求每个访问的应用都有代理。第三，在这种防火墙中，一旦一个连接建立起来，就不用再对这个连接做更多工作，系统可以去处理别的连接。这些都使状态检测防火墙执行效率明显提高。

3）更好的扩展性

状态检测防火墙不像应用网关式防火墙那样，每一个应用对应一个服务程序，这样所能提供的服务是有限的，而且当增加一个新的服务时，必须为新的服务开发相应的服务程序，这样系统的可扩展性降低。状态检测防火墙不区分每个具体的应用，只是根据从数据包中提取出的信息、对应的安全策略及过滤规则处理数据包，当有一个新的应用时，它能动态产生新的应用规则，而不用另外写代码，所以具有很好的伸缩性和扩展性。

4）配置方便，应用范围广

状态检测防火墙不仅支持基于 TCP 的应用，而且支持基于无连接协议的应用，如 RPC、

基于 UDP 的应用（DNS、WAIS、Archie 等）等。对于无连接的协议，连接请求和应答没有区别，包过滤防火墙和应用网关对此类应用要么不支持，要么开放一个大范围的 UDP 端口，这样暴露了内部网，降低了安全性。而这样的攻击都可以被状态检测防火墙阻塞，它通过控制无效连接的连接时间，避免大量的无效连接占用过多的网络资源，可以很好地降低 DoS 和 DDoS 攻击的风险。

状态检测防火墙也支持 RPC，因为对于 RPC 服务来说，其端口号是不定的，因此简单地跟踪端口号不能实现该种服务的安全，状态检测防火墙通过动态端口映射图记录端口号，为验证该连接，还保存连接状态、程序号等，通过动态端口映射图来实现此类应用的安全。

#### 4. 状态检测防火墙的缺点

状态检测防火墙虽然继承了包过滤防火墙和应用网关防火墙的优点，克服了它们的缺点，但它仍只是检测数据包的第三层信息，无法彻底识别数据包中大量的垃圾邮件、广告以及木马程序等。

包过滤防火墙、网关代理防火墙以及状态检测防火墙都有固有的无法克服的缺陷，不能满足用户对于安全性的不断的要求，于是深度包检测防火墙技术被提出了。

### 4.1.6 网络防火墙部署

#### 1. 屏蔽路由器（screening router）和屏蔽主机（screening host）

防火墙最基本，也是最简单的技术是数据包过滤。过滤规则可以安装在路由器上，也可以安装在主机上。具有数据包过滤功能的路由器称为屏蔽路由器，具有数据包过滤功能的主机称为屏蔽主机。图 4.9 为屏蔽路由防火墙的基本结构。

路由器是内部网络与 Internet 连接的必要设备，是一种“天然”的防火墙，它除具有路由功能之外，还安装了分组/包过滤（数据包过滤或应用网关）软件，可以决定对到来的数据包是否要进行转发。这种防火墙实现方式简捷，效率较高，在应用环境比较简单的情况下，能够以较小的代价在一定程度上保证系统的安全。但它也有许多不足。

（1）过滤路由器是在网关之上实施包过滤，因此它允许被保护网络的多台主机与 Internet 的多台主机直接通信。这样，其危险性便分布在被保护网络内的全部主机以及允许访问的各种服务器上；服务越多，网络的危险性也就越大。

（2）这种网络仅靠单一的部件来保护系统，一旦部件被攻破，就再也没有任何设防了，并且防火墙被攻破时几乎不留下任何痕迹，难以发现已发生的攻击。

（3）它只能根据数据包的源/目的地址和端口等网络信息进行判断，无法识别基于应用层的恶意侵入，如恶意的 Java 小程序以及电子邮件中附带的病毒。有经验的黑客很容易伪造 IP 地址骗过包过滤型防火墙，直接对主机上的软件和配置漏洞进行攻击。

（4）由于数据包的源地址、目的地址以及 IP 的端口号都在数据包的头部，很有可能被窃听或假冒。

（5）数据包缺乏用户日志（log）和审计信息（audit），不具备登录和报告性能，不能进

行审核管理，因而过滤规则的完整性难以验证，所以安全性较差。

**2. 双宿主主机**

如图 4.10 所示，双宿主主机（dual homed host，也称为双穴主机）是至少有两个网络接口的主机，每个接口有一块 NIC（Network Interface Card，网络接口卡，简称网卡，也称为网络适配器），各有一个 IP 地址。它具有如下功能。

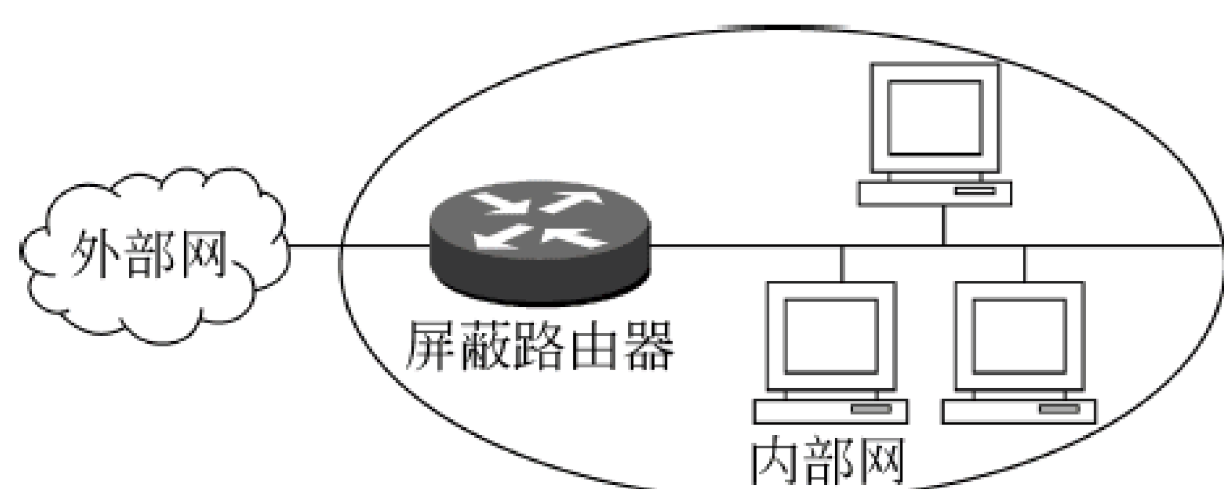



图 4.9 屏蔽路由防火墙

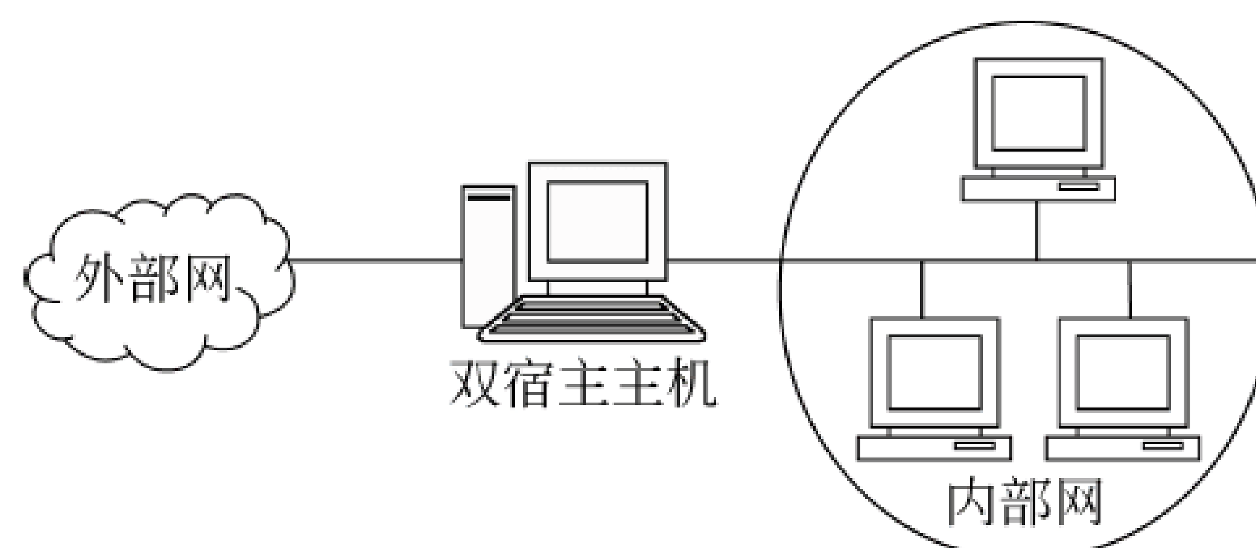



图 4.10 双宿主主机网关防火墙

（1）在所连接的多个网络之间建立一个阻塞点，从一个网络到另一个网络发送的 IP 数据包必须经过双宿主主机的检查。

（2）与它相连的内部和外部网络都可以执行由它所提供的网络应用，如果这个应用允许的话，它们就可以共享数据。

这样就保证内部网络和外部网络的某些节点之间可以通过双宿主主机上的共享数据传递信息，但内部网络与外部网络之间却不能直接传递信息，从而达到保护内部网络的作用。

在双宿主主机中可以采用 NAT 和代理两种安全机制。如果 Internet 上的一台计算机想与被保护网络（Intranet）上的一个工作站通信，必须先注册，与它能看到的 IP 地址联系；代理服务器软件通过另一块 NIC 启动到 Intranet 的连接。

双宿主主机使用代理服务器简化了用户的访问过程，它将被保护网络与外界完全隔离，由于域名系统的信息不会通过被保护系统传到外部，所以系统的名字和 IP 地址对 Internet 是隐蔽的，做到对用户全透明。由于该防火墙仍是由单机组成的，没有安全冗余机制，一旦该“单失效点”出问题，网络将无安全可言。

**3. 堡垒主机**

1）堡垒主机（bastion host）的特性

（1）它是专门暴露给外部网络上的一台计算机，是被保护的内部网络在外网上的代表，并作为进入内部网的一个检查点。

（2）它面对大量恶意攻击的风险，并且它的安全对于建立一个安全周边具有重要作用，因此必须强化对它的保护，使风险降至最小。

（3）它通常提供公共服务，如邮件服务、WWW 服务、FTP 服务和 DNS 服务等。

（4）堡垒主机与内部网络是隔离的。它不知道内部网络上的其他主机的任何系统细节，如内部主机的身份认证服务和正在运行的程序的细节。这样，对堡垒主机的攻击不会殃及内部网络。

所以，堡垒主机是一个被强化的、被暴露在被保护网络外部的、可以预防进攻的计

算机。

2）堡垒主机的单连点和双连点结构

（1）单连点堡垒主机过滤式防火墙。单连点堡垒主机过滤式防火墙的结构如图 4.11 所示。它实现了网络层安全（包过滤）和应用层安全（代理），具有比单纯包过滤更高的安全等级。

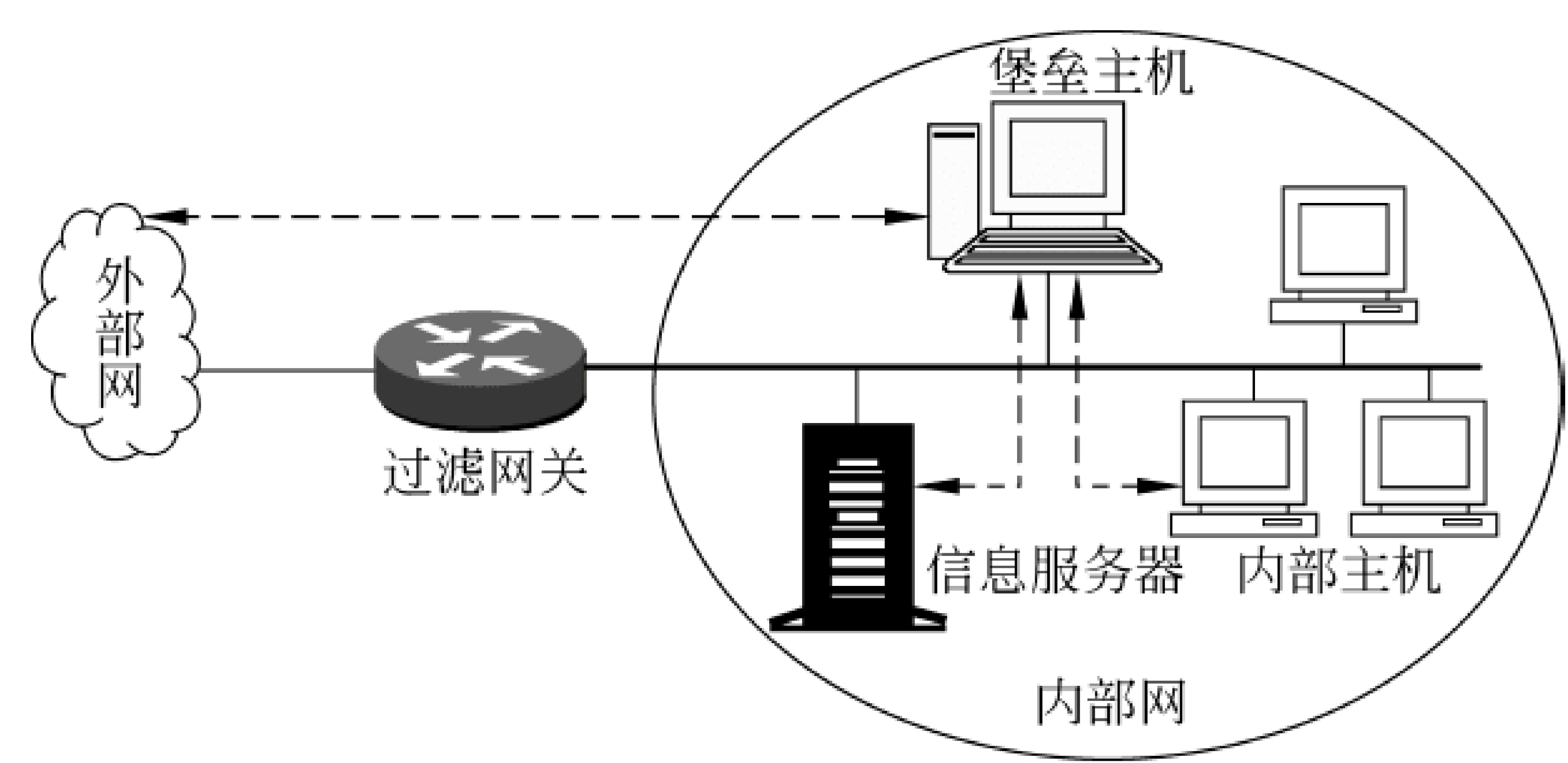



图 4.11　单连点堡垒主机过滤式防火墙

在该系统中，堡垒主机被配置在过滤路由器的后方，并且过滤规则的配置使得外部主机只能访问堡垒主机，发往内部网的其他业务流则全部被阻塞。对于内部主机来说，由于内部主机和堡垒主机同在一个内部网络上，所以机构的安全策略可以做出以下决定：是内部系统允许直接访问外部网，还是要求使用配置在堡垒主机上的代理服务。当配置路由器的过滤规则使其仅仅接收来自堡垒主机的内部业务流时，内部用户就不得不使用代理服务。

主机过滤防火墙具有双重保护，从外网来的访问只能到达堡垒主机，而不允许访问被保护网络的其他资源，有较高的安全可靠性。并且主机过滤网关能有选择地允许那些可以信赖的应用程序通过路由器，是一种非常灵活的防火墙。但是它要求考虑到堡垒主机和路由器两个方面的安全性。如果路由器中的访问控制表允许某些访问通过路由器，则防火墙管理员不仅要管理堡垒主机中的访问控制表，而且要管理路由器中的访问控制表，并要求对这两个部件仔细配置以便它们能协调工作。此外，系统的灵活性也会导致走捷径（例如用户可能试图避开代理服务器直接与路由器建立联系）而破坏安全性。

（2）双连点堡垒主机过滤式防火墙。

双连点堡垒主机过滤式防火墙的结构如图 4.12 所示。它比单连点堡垒主机过滤式防火

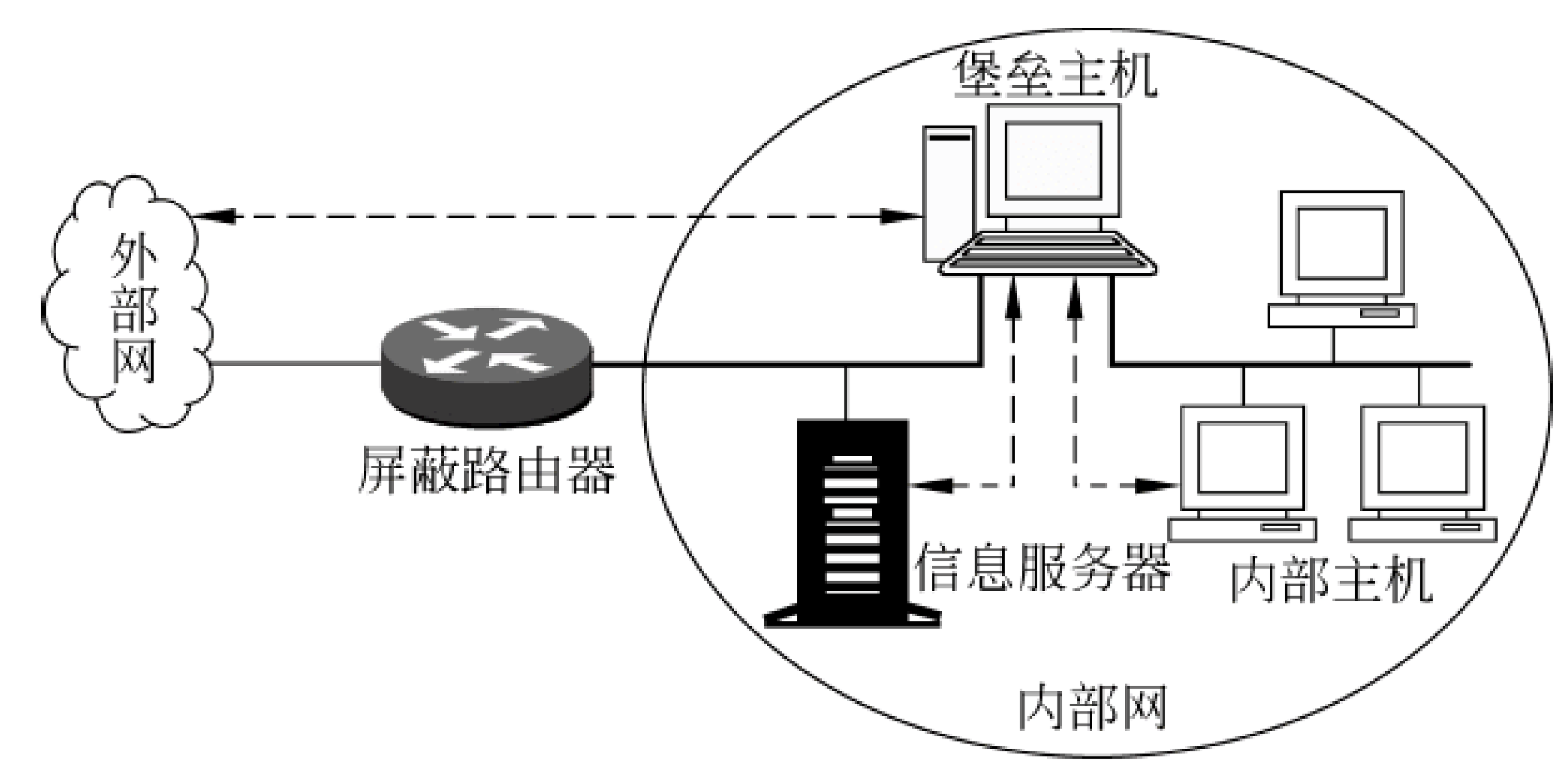



图 4.12　双连点堡垒主机过滤式防火墙

墙有更高的安全等级。由于堡垒主机具有两个网络接口，除了外部用户可以直接访问信息服务器外，外部用户发往内部网络的业务流和内部系统对外部网络的访问都不得不经过堡垒主机，以提高附加的安全性。

在这种系统中，由于堡垒主机成为外部网络访问内部网络的唯一入口，所以对内部网络的可能安全威胁都集中到了堡垒主机上。因而对堡垒主机的保护强度关系到整个内部网的安全。

#### 4. 屏蔽子网防火墙

屏蔽子网（screened subnet）防火墙是在被保护网络和 Internet 之间设置一个独立的子网作为防火墙。具体的配置方法是在过滤主机的配置上再加上一个路由器，形成具有外部路由过滤器、内部路由过滤器和应用网关 3 道防线的过滤子网，如图 4.13 所示。

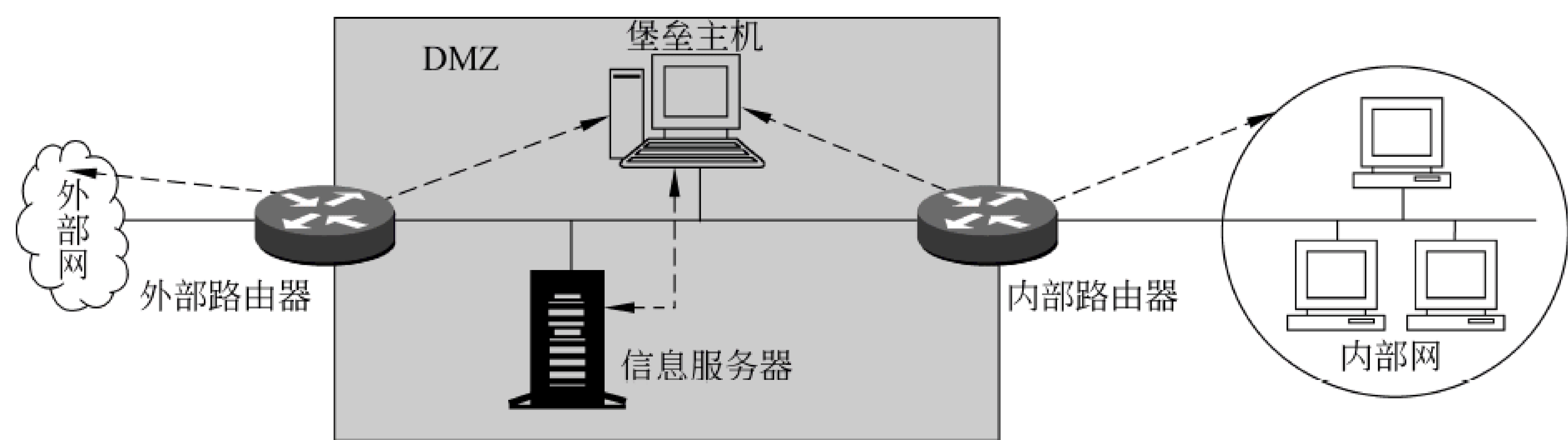



图 4.13　子网过滤防火墙配置

在屏蔽子网防火墙中，外部过滤路由器用于防范通常的外部攻击（如源地址欺骗和源路由攻击），并管理外部网到过滤子网的访问。外部系统只能访问到堡垒主机，通过堡垒主机向内部网络传送数据包。内部过滤路由器管理过滤子网与内部网络之间的访问，内部系统也只能访问到堡垒主机，通过堡垒主机向外部网络发送数据包。简单地说，任何跨越子网的直接访问都是被严格禁止的，从而在两个路由器之间定义了一个“非军事区”（De-Militarized Zone，DMZ），表明内部网络距离外部网络更远，更安全。这种配置的防火墙具有最高的安全性，但是它要求的设备和软件模块较多，价格较贵，且相当复杂。

## 4.2　网络的物理隔离技术

### 4.2.1　物理隔离的概念

#### 1. 问题的提出

20 世纪末期，“政府上网”热潮把我国的信息化带进了一个新的高度。政府上网不仅表明 Internet 已经进入了一个非常重要的领域，而且为信息系统安全技术提出了新的课题。政府中有许多敏感的数据以及大量机密数据，也有最为国民关心的信息。目前虽然开发了各种防火墙、病毒防治系统以及入侵检测、安全预警、漏洞扫描等安全技术，但却并没有完全阻止入侵，内部网络被攻破的事件屡有发生，据统计有近半数的防火墙被攻破过。为此，国家保密局 2000 年 1 月 1 日起实施的《计算机信息系统国际联网保密管理规定》第二章第

六条要求："涉及国家机密的计算机信息系统，不得直接或间接地与国际互联网或其他公共信息网络相连接，必须实行物理隔离。"

学术界一般认为，最早提出物理隔离技术的是以色列和美国的军方，主要用于解决涉密网络与公共网络连接时的安全。我国也有庞大的政府涉密网络和军事涉密网络，但是我国的涉密网络与公共网络，特别是与 Internet 无任何关联的独立网络，不存在与 Internet 的信息交换，也用不着使用物理隔离网闸解决信息安全问题。所以，在电子政务、电子商务之前，物理隔离网闸在我国因无市场需求，产品和技术发展较慢。

随着我国信息化建设步伐的加快，电子政务的急速展开，物理隔离的问题才突出地提到议事日程上来，成为我国信息安全产业发展的一个新的增长点。2002 年 8 月 15 日，《国家信息化领导小组关于我国电子政务建设的指导意见》（中办 17 号文件）提出了"十五"期间我国电子政务建设的主要任务之一是："建设和整合统一的电子政务网络。为适应业务发展和安全保密的要求，有效遏制重复建设，要加快建设和整合统一的网络平台。电子政务网络由政务内网和政务外网构成，两网之间物理隔离，政务外网与 Internet 之间逻辑隔离。政务内网主要是副省级以上政务部门的办公网，与副省级以下政务部门的办公网物理隔离。政务外网是政府的业务专网，主要运行政务部门面向社会的专业性服务业务和不需在内网上运行的业务。要统一标准，利用统一平台，促进各个业务系统的互联互通、资源共享。要用一年左右的时间，基本形成统一的电子政务内外网络平台，在运行中逐步完善。"

简单地说，如图 4.14 所示，政务网应当跨越公网、外网和内网。其安全要求如下。

（1）在公网和外网之间实行逻辑隔离。

（2）在内网和外网之间实行物理隔离。

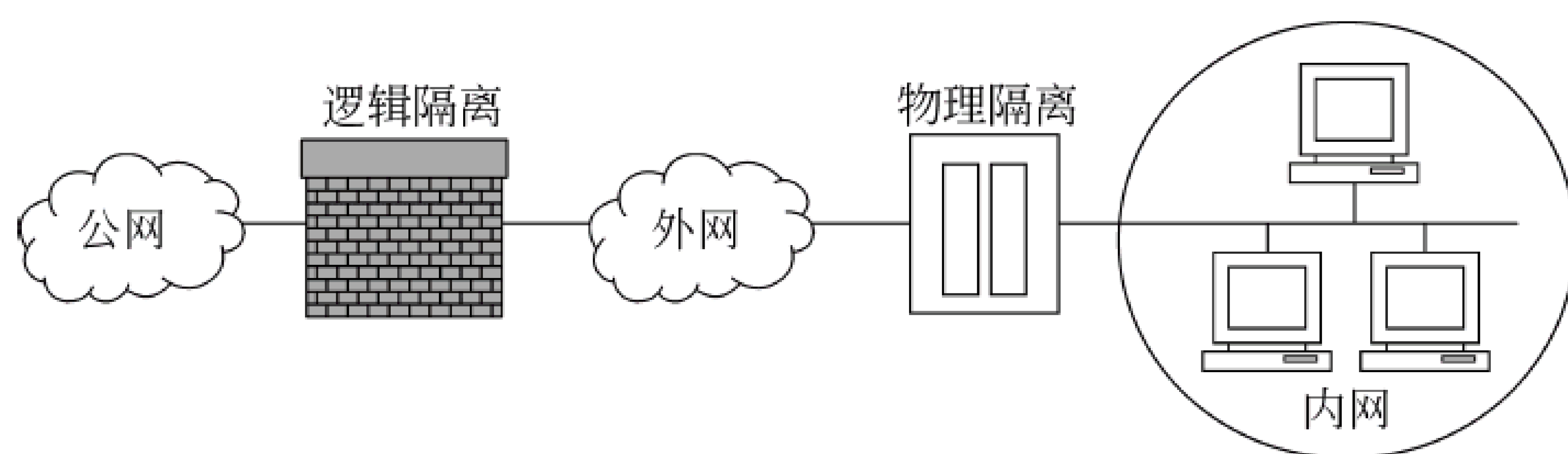



图 4.14　电子政务的三网

**2. 物理隔离的术语及其理解**

但是到目前为止，并没有完整的关于物理隔离技术的定义和标准。从不同时期的用词也可以看出，物理隔离技术一直在演变和发展。

较早的用词为 Physical Disconnection。Disconnection 有使断开、切断、不连接的意思，直译为物理断开。这是在还没有解决涉密网与 Internet 连接后出现的很多安全问题的技术手段之前的说法，在无可奈何的情况下，只有先断开再说。

后来使用了词汇 Physical Separation。Separation 有分开、分离间隔和距离的意思，直译为物理分开。

但是光分开不是办法，理智的策略应当是为该连即连，不该连则不连。为此要把该连的部分与不该连的部分分开。于是有了 Physical Isolation。Isolation 有孤立、隔离、封闭、绝缘的意思，直译为物理封闭。

事实上，没有与 Internet 相联的系统不多，因此，希望能将一部分高安全性的网络隔离封闭起来。于是开始使用 Physical Gap。Gap 有豁口、裂口、缺口和差异的意思，直译为物理隔离，意为通过制造物理的豁口来达到隔离的目的。

由于 Physical 这个词显得非常僵硬，于是有人用 Air Gap 来代替 Physical Gap。Air Gap 意为空气豁口，很明显在物理上是隔开的。但有人不同意，理由是空气豁口就“物理隔离”了吗？电磁辐射、无线网络、卫星等都是空气豁口，却没有物理隔离，甚至连逻辑上都没有隔离。于是，E-Gap、Netgap、I-Gap 等都出来了。现在，一般称为 Gap Technology，意为物理隔离，成为 Internet 上的一个专用名词。

### 3. 对物理隔离的要求及其理解

物理隔离要求内部网络与外部网络在物理上没有相互连接的通道，两个系统在物理上完全独立。要实现公众信息网（外部网）与内部网络物理隔离的目的，必须保证做到以下几点。

（1）在物理传导上使内外网络隔断，确保外部网不能通过网络连接而侵入内部网；同时防止内部网信息通过网络连接泄露到外部网。

（2）在物理辐射上隔断内部网与外部网，确保内部网信息不会通过电磁辐射或耦合方式泄露到外部网。

（3）在物理存储上隔断两个网络环境，对于断电后会易失信息的部件，如内存、处理器等暂存部件，要在网络转换时进行清除处理，防止残留信息串网；对于断电非易失性设备，如磁带机、硬盘等存储设备，内部网与外部网信息要分开存储。

具体地说，对物理隔离的理解表现为以下几个方面。

① 阻断网络的直接连接，即没有两个网络同时连在隔离设备上。

② 阻断网络的 Internet 逻辑连接，即 TCP/IP 的协议必须被剥离，将原始数据通过 P2P 的非 TCP/IP 连接协议透过隔离设备传递。

③ 隔离设备的传输机制具有不可编程的特性，因此不具有感染的特性。

④ 任何数据都通过两级移动代理的方式来完成，两级移动代理之间是物理隔离的。

⑤ 隔离设备具有审查的功能。

⑥ 隔离设备传输的原始数据不具有攻击或对网络安全有害的特性。就像 txt 文本不会有病毒也不会执行命令等一样。

⑦ 强大的管理和控制功能。

### 4. 数据隔离与网络隔离

从隔离的内容看，隔离分为数据隔离和网络隔离。

（1）数据隔离主要是指存储设备的隔离——一个存储设备不能被几个网络共享。

（2）网络隔离就是把被保护的网络从公开的、无边界的、自由的环境中独立出来。

只有实现了上述两种隔离，才是真正意义上的物理隔离。

**5. 逻辑隔离部件与物理隔离部件**

物理隔离与逻辑隔离有很大的不同。物理隔离的哲学是不安全就不联网，要绝对保证安全；逻辑隔离的哲学是在保证网络正常使用的情况下尽可能安全。在技术上，实现逻辑隔离的方式有很多，但主要是防火墙。

中华人民共和国公安部 2001 年 12 月 24 日发布（2002 年 5 月 1 日实施）的《端设备部件安全技术要求》（GA 370—2001）指出：

（1）物理隔离部件的安全功能应保证被隔离的计算机资源不能被访问（至少应包括硬盘、软盘和光盘），计算机数据不能被重用（至少应包括内存）。

（2）逻辑隔离部件的安全功能应保证被隔离的计算机资源不能被访问，只能进行隔离器内外的原始应用数据交换。

（3）单向隔离部件的安全功能应保证被隔离的计算机资源不能被访问（至少应包括硬盘/硬盘分区、软盘和光盘），计算机数据不能被重用（至少应包括内存）。

（4）逻辑隔离部件应保证其存在泄露网络资源的风险不得多于开发商的评估文档中所提及的内容。

（5）逻辑隔离部件的安全功能应保证在进行数据交换时数据的完整性。

（6）逻辑隔离部件的安全功能应保证隔离措施的可控性，隔离的安全策略应由用户进行控制，开发者必须提供可控方法。

（7）单向隔离部件使数据流无法从专网流向外网，数据流能在指定存储区域从公网流向专网；对专网而言，还能使用外网的某些指定的导入数据。

图 4.15 为使用物理隔离部件和单向隔离部件进行连接的示意图。

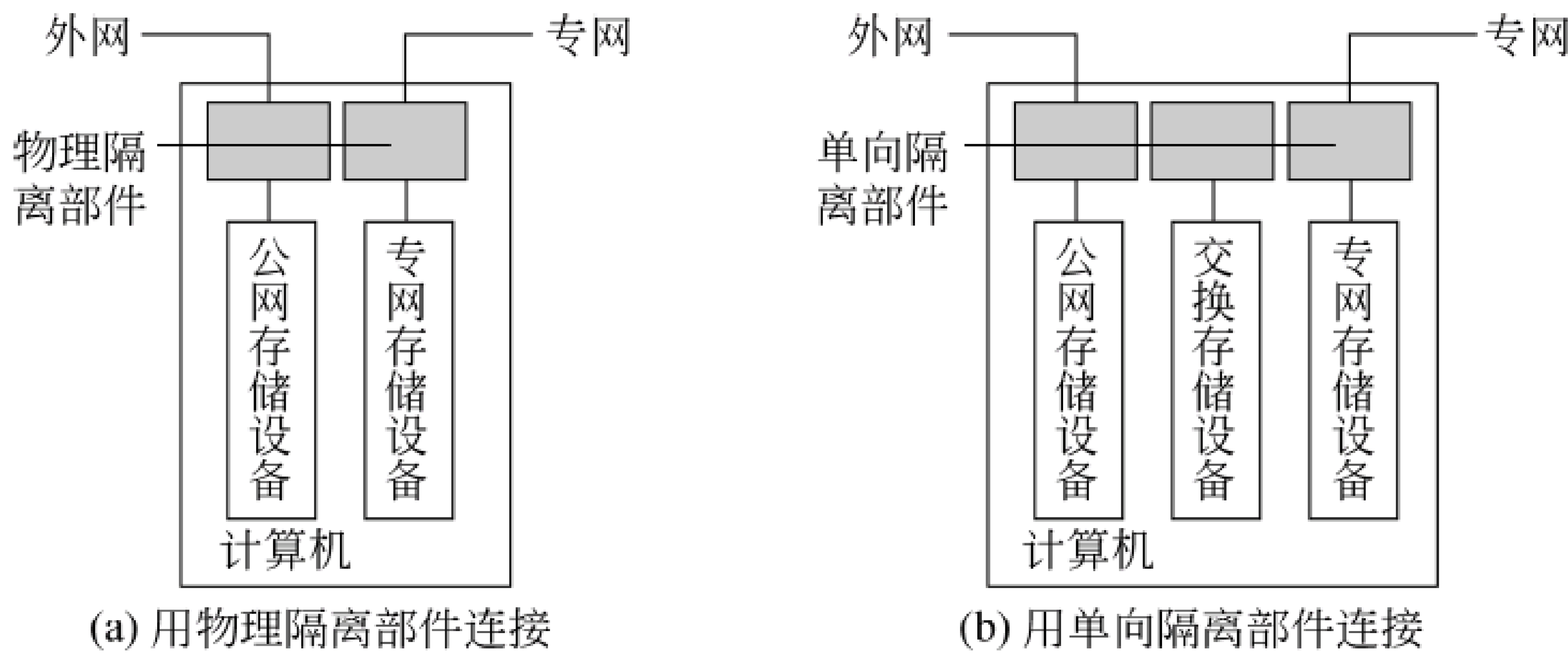



图 4.15　物理隔离部件和单向隔离部件的连接示意图

目前物理隔离技术主要包括网络安全隔离卡、隔离集线器和网闸 3 种。

### 4.2.2　网络安全隔离卡

网络安全隔离卡是一种用户级物理隔离技术，其基本原理是在计算机中使用两个独立的硬盘，或将一个硬盘分割成两个独立的物理区。

（1）安全区，只与内部网络连接。

（2）公共区，只与外部网络连接。

如图 4.16 所示，网络安全隔离卡就像一个分接开关，在 IDE 硬件层上，由固件控制磁

盘通道，任何时刻计算机只能与一个数据分区以及相应的网络连通。于是计算机也因此被分为安全状态和公共状态，并且某一时刻只可以在一个状态下工作。

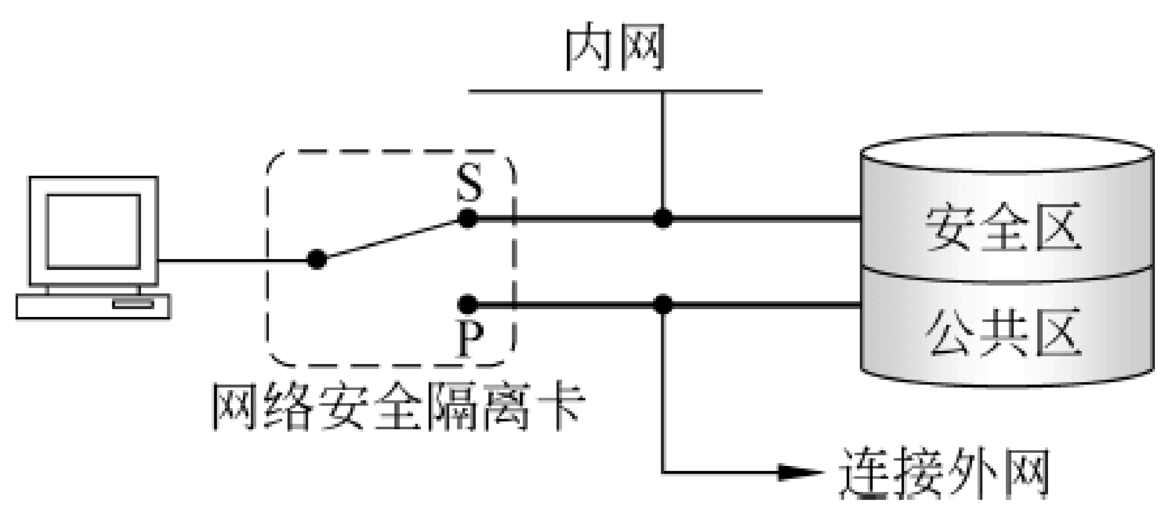


图 4.16　网络安全隔离卡的工作方式

（1）在安全状态时，主机只能使用硬盘的安全区与内网连接，此时外网是断开的，硬盘的公共区也是封闭的。

（2）在公共状态时，主机只能使用硬盘的公共区与外网连接，此时与内网是断开的，且硬盘的安全区是封闭的。

两个状态各有自己独立的操作系统，并分别导入，保证两个硬盘不会同时被激活。两个分区不可以直接交换数据，但是可以通过专门设置的中间功能区进行，或通过设置的安全通道使数据由公共区向安全区转移（不可逆向）。两个状态转换时，所有的临时数据都会被彻底删除。

在安全区及内网连接状态下可以禁用软驱、光驱等移动存储设备，防止内部数据泄密。要转换到公共环境时，须进行如下操作。

① 按正常方式退出操作系统。

② 关闭计算机。

③ 将安全硬盘转换为公共硬盘。

④ 将 S/P 开关转换到公共网络。

当然，这些操作是由网络安全隔离卡自动完成的。为了便于用户从 Internet 上下载数据，特设硬盘数据交换区，通过读写控制只允许数据从外网分区向内网分区单向流动。

图 4.17 为一种客户端物理隔离系统的解决方案。

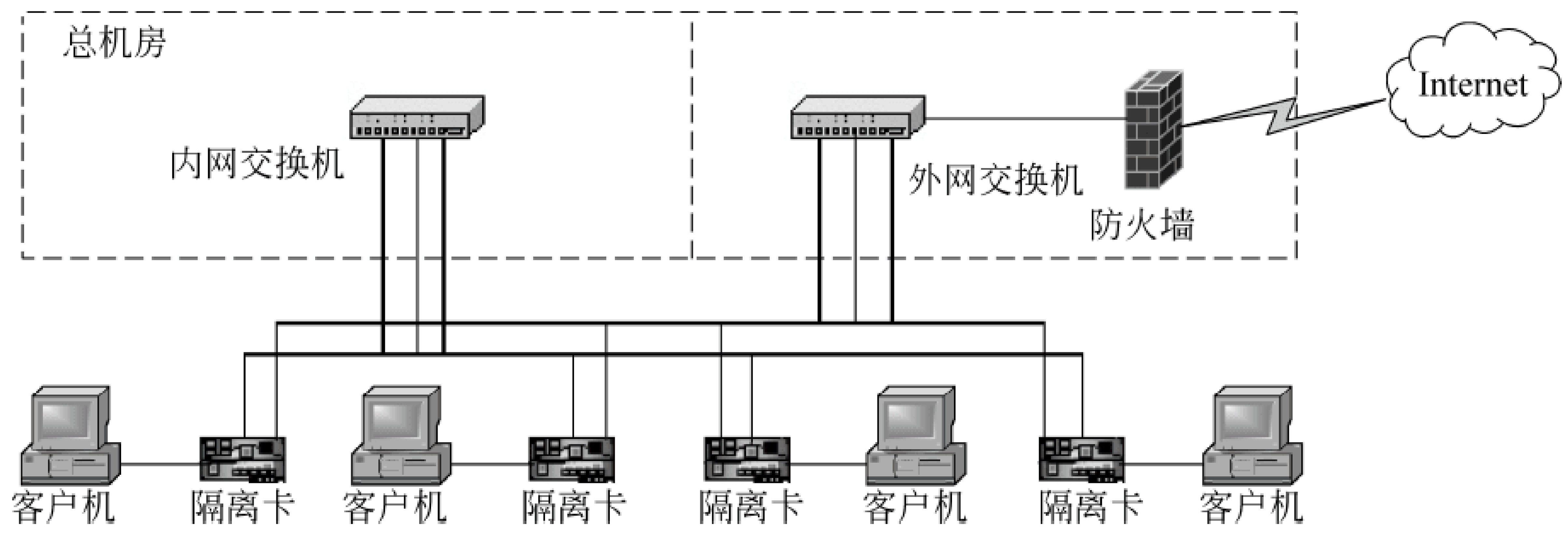


图 4.17　一种客户端物理隔离系统解决方案

### 4.2.3　隔离集线器技术

如图 4.18 所示，网络安全集线器是一种多路开关切换设备。它与网络安全隔离卡配合使用，并通过对网络安全隔离卡上发出的特殊信号的检测，识别出所连接的计算机，自动将其网线切换到相应的网络的 Hub 上，从而实现多台独立的安全计算机与内、外两个网络的安全连接与自动切换。

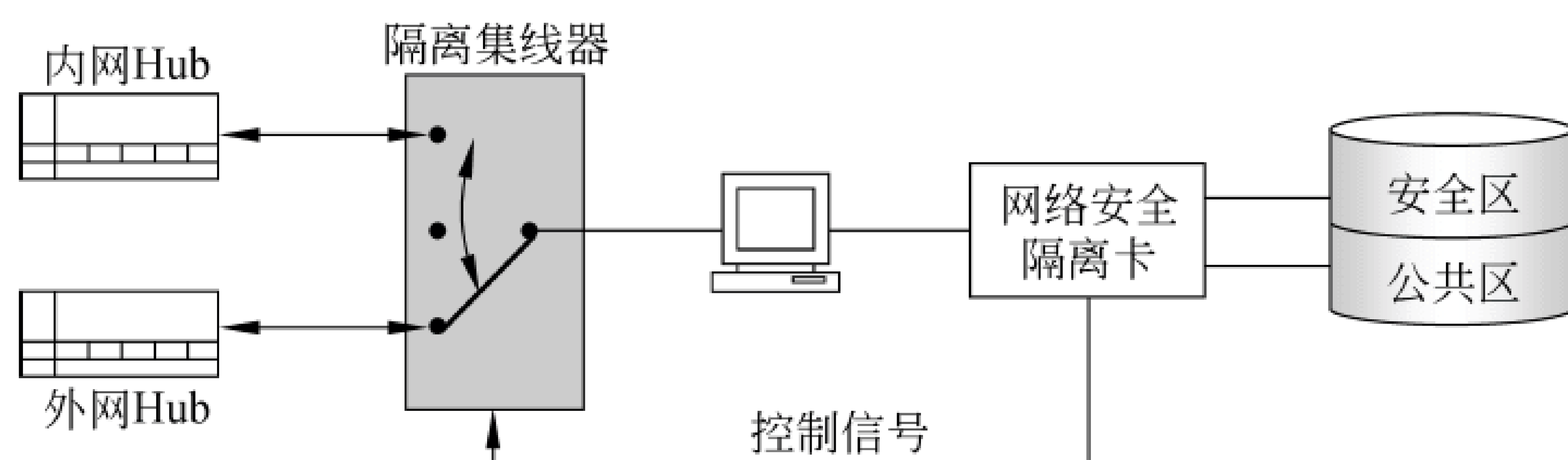


图 4.18　网络安全隔离集线器的工作原理

如果没有检测到来自网络安全隔离卡的信号，两个网络都会被切断。这样减少了安全区的工作站被错误地连入未分类网络的风险。

**注意：**隔离集线器只有与其他隔离措施，如物理隔离卡等相配合，才能实现真正的物理隔离。如果只切换内外网且变更 IP 地址，而不重新启动系统，则不是真正的物理隔离。图 4.19 为一种采用隔离集线器和物理隔离卡的解决方案。

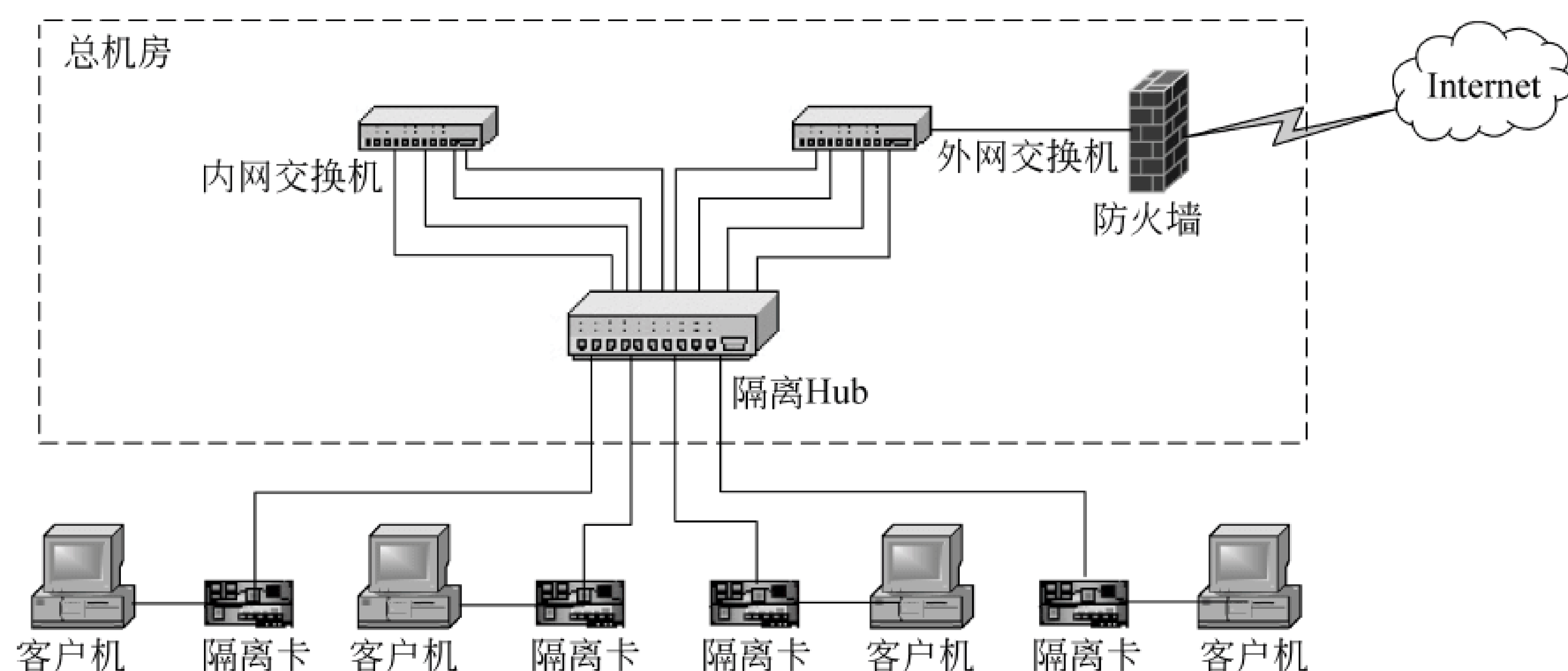


图 4.19　一种采用隔离集线器和物理隔离卡的解决方案

在隔离集线器的基础上，有人提出了隔离交换机的方案。

## 4.2.4　网闸

### 1. 网闸的基本原理

计算机网络是基于网络协议实现连接的。几乎所有的攻击都是在网络协议的一层或多层上进行的。理论上讲，如果断开 OSI 数据模型的所有层，就可以消除来自网络的潜在攻击。网闸正是依照此原理实现了信息安全传递。

网闸的全称是安全隔离网闸，也称为信息交换与安全隔离系统（官方称呼），其设计理念基于如下两点。

（1）“摆渡”。

过河可以用船摆渡，也可以通过桥。差别在于摆渡不连接两岸，桥连接两岸。网闸采用摆渡方式。如图 4.20 所示，用这种方式进行网络隔离，即当设备连接一端时，另一端一定是断开的，这就断开了物理层和数据链路层，消除了物理层和数据链路层的漏洞。

（2）“裸体检查”。

传统网络传输是经过一层一层地封装，在每一层中按照协议进行转发。这些转发可以

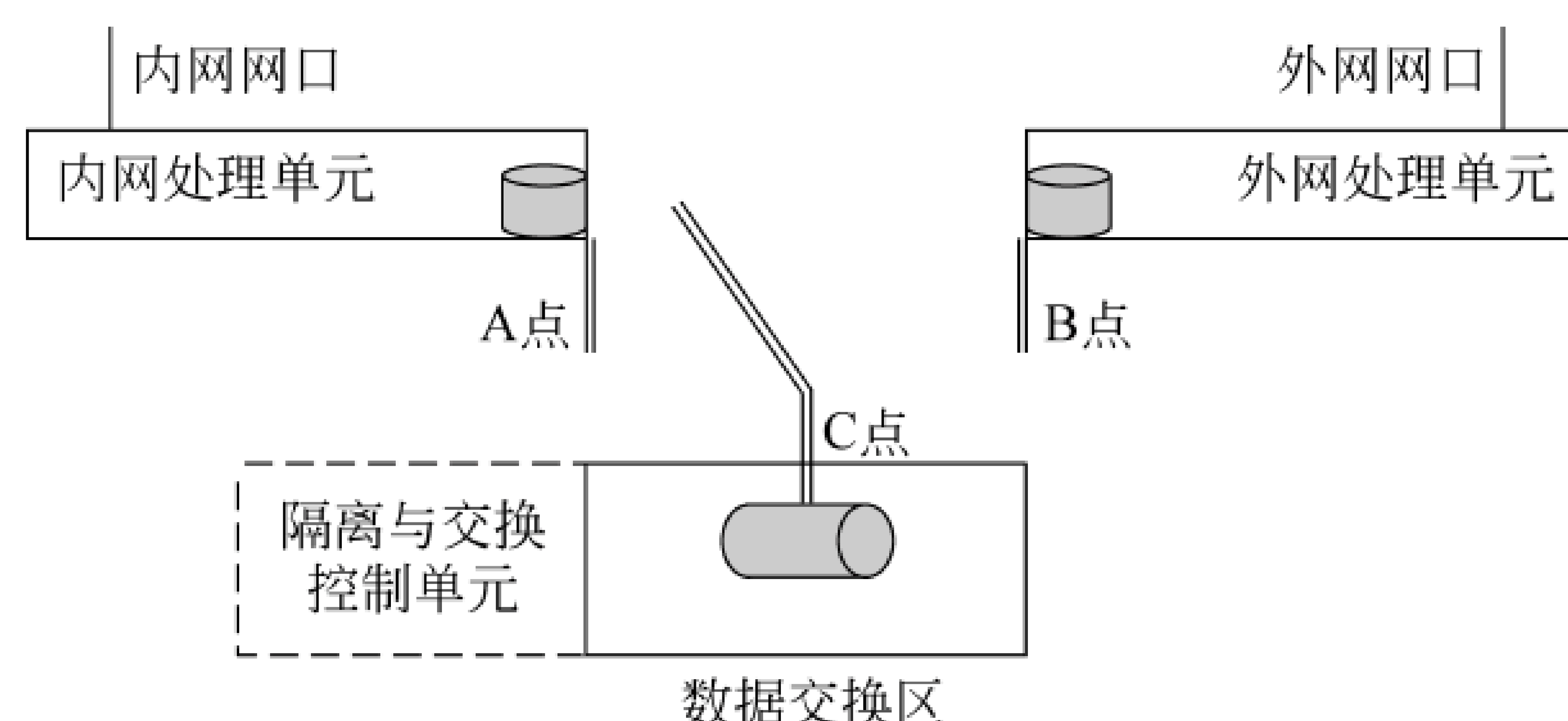



图 4.20　网络“摆渡”示意图

称为逻辑连接。摆渡消除了物理层和数据链路层的漏洞，但无法消除之上各层的漏洞。要想通过摆渡网闸消除 TCP/IP（OSI 的 3 到 4 层）漏洞，必须剥离 TCP/IP；要想消除应用协议（OSI 的 5 到 7 层）漏洞，必须剥离应用协议。这是一种“裸体检查的”思想，好像体检要脱光衣服一样。在网闸工作时，经过了一个剥离—检测—重封装的过程，即首先将数据包进行包头的剥离、分解或重组，然后对静态的裸数据进行安全审查（包括网络协议检查和代码扫描等），之后用特殊的内部协议封装后转发，到达对端网络后再重新按照 TCP/IP 进行封装，实现了“协议落地，内容检测”。

图 4.21 是一个采用网闸的政府网络安全解决方案。

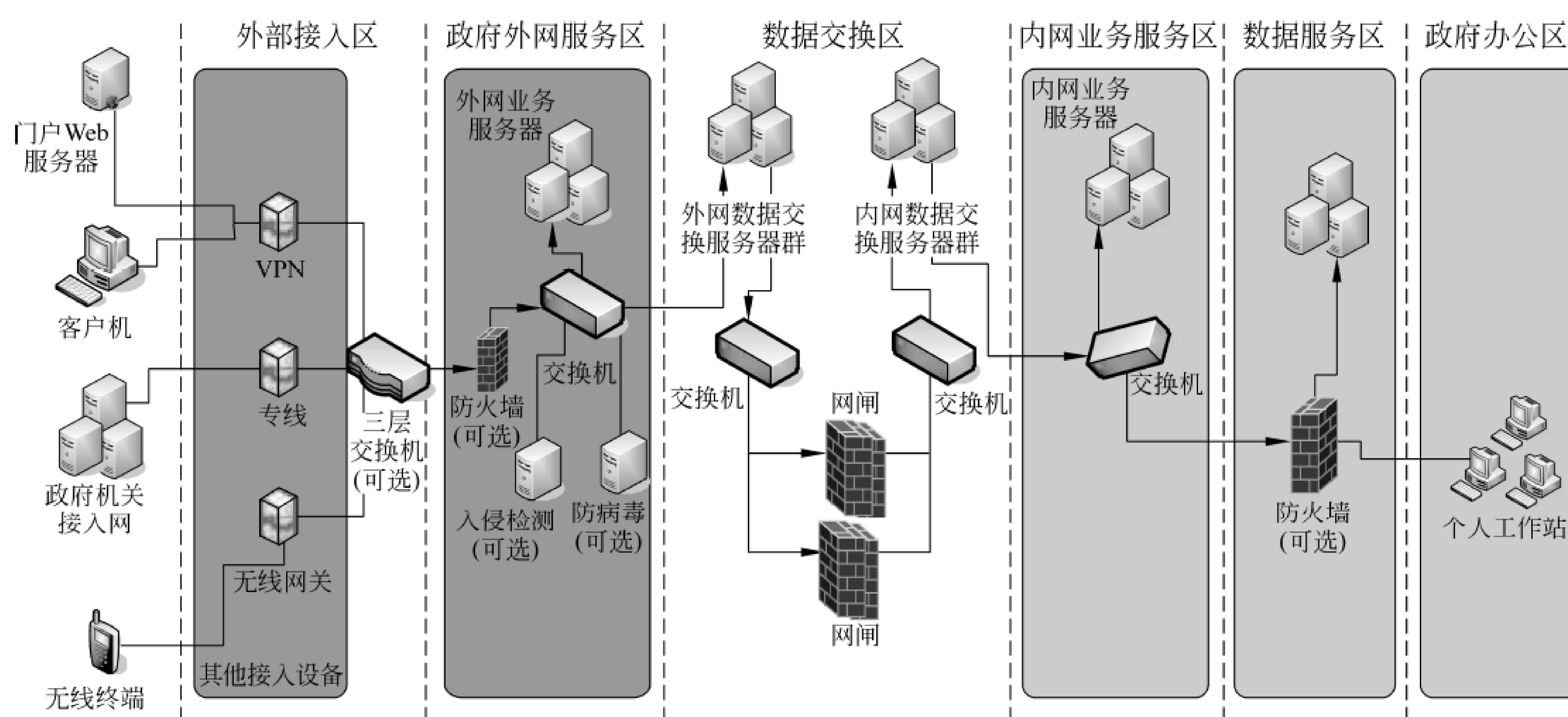



图 4.21　一个采用网闸的政府网络安全解决方案

### 2. 网闸的基本结构

如图 4.22 所示，网闸系统主要由内网处理模块、外网处理模块和数据交换模块 3 个模块组成。其中，内、外网处理模块负责内、外网信息获取和协议分析；而安全检测与控制处理模块则根据安全策略完成信息的安全检测、内外网络隔离和安全交换，有的还具有更完善的功能，如内核防护、协议转换、病毒查杀、访问控制、安全审计、身份认证等。

网闸的主要性能指标有系统数据交换速率（一般要求大于 120Mbps）和硬件切换时间（一般要求小于 5ms）。

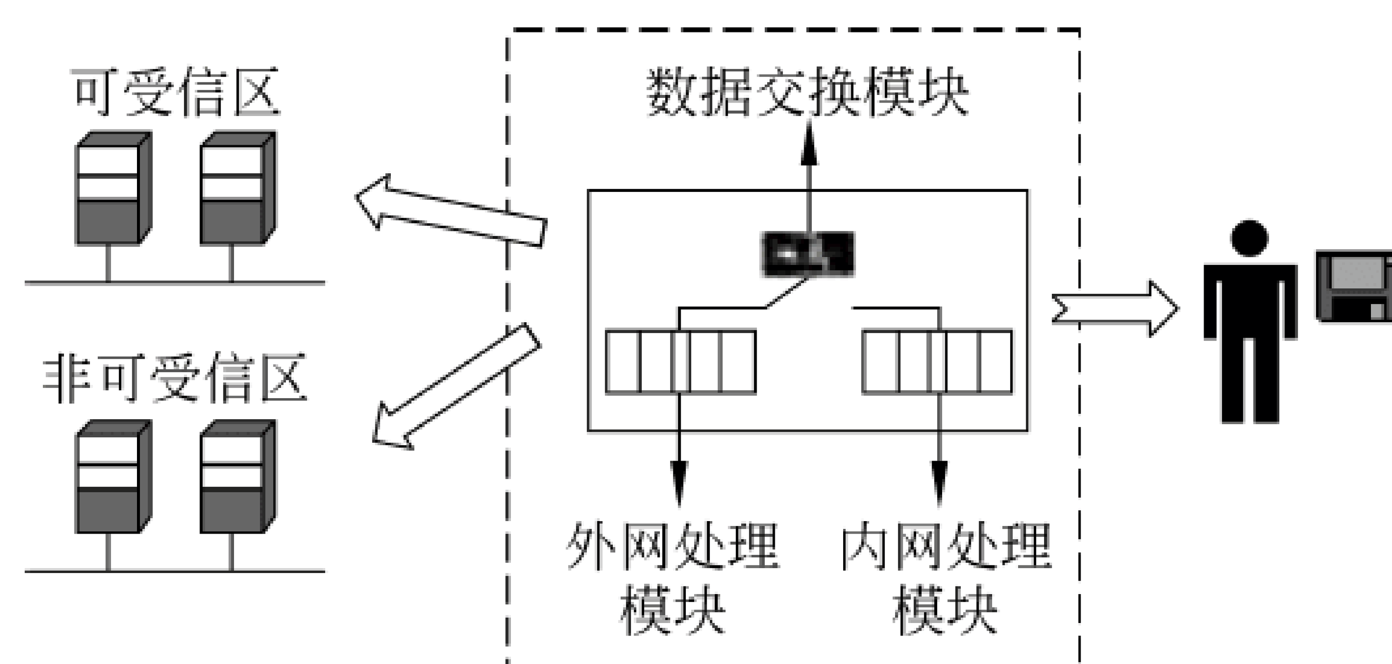



图 4.22　网闸的基本结构

### 3. 网闸的实现技术

下面介绍目前的网闸三大实现技术。

1）基于存储总线的网闸技术

基于存储总线的网闸技术是目前最主流的网闸技术。图 4.23 为基于存储总线的网闸工作原理示意图，它用可读写本地存储介质作为交换区，利用电子开关控制内外网的主机单元以摆渡方式进行数据交换区内存储空间的读写。主机单元通过扩展卡扩展存储总线，并把由控制信号组成的专用扩展总线连接到隔离交换控制主机上。对交换区的读写采取块方式，不能采用文件方式，数据的校验和文件的还原都在主机单元中进行；需要读出写入的数据，通过比较来确认写入的数据是否正确。

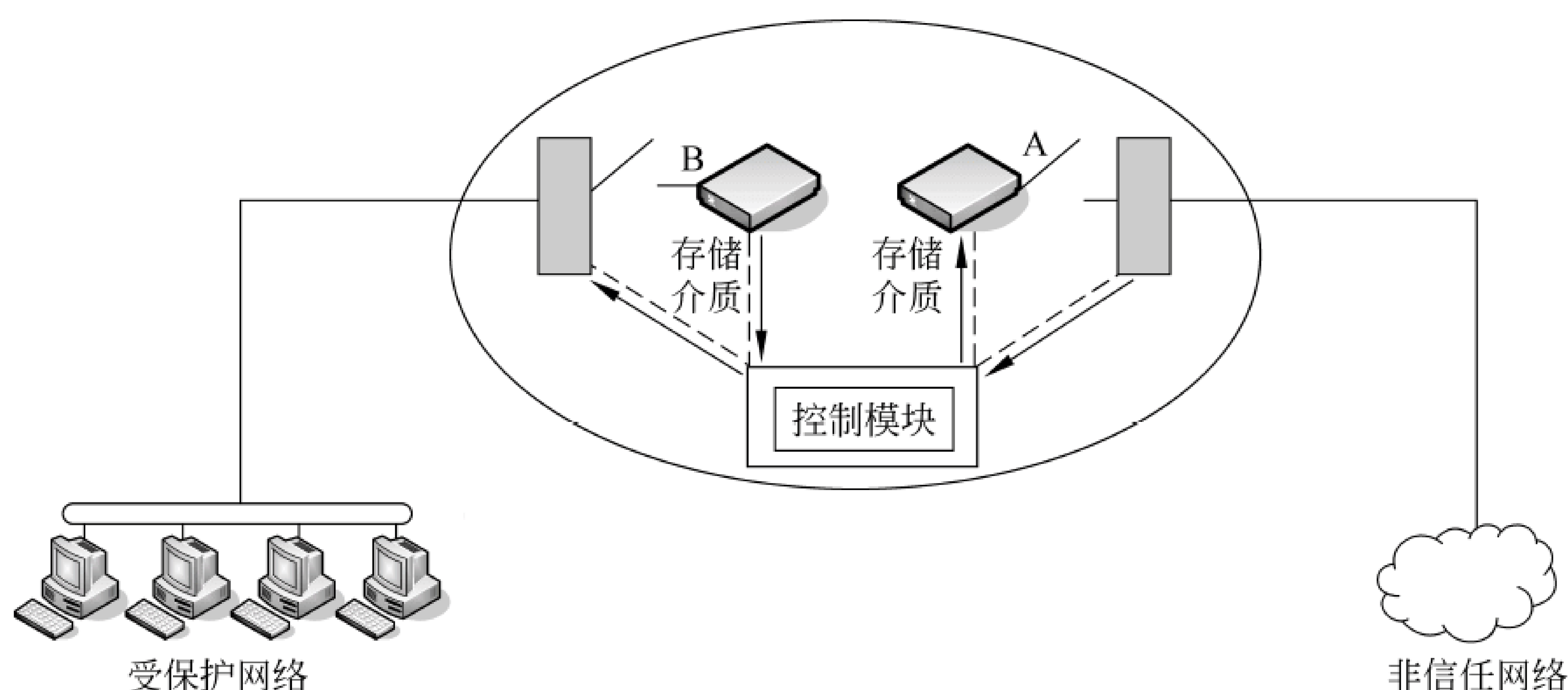



图 4.23　一种基于存储总线的网闸结构

具体的存储技术可以采用 SCSI 方式，也可以采用 IDE 方式。从设计的方便上还可以选择串行的存储方式，如 SATA、SAS 或 USB 盘方式。

2）基于通信总线的网闸技术

基于通信总线的网闸技术也是目前成熟的技术之一。如图 4.24 所示，这种技术采用双端口的静态存储器（Dual Port SRAM），配合基于独立的 CPLD 的控制电路，以实现在两个端口上的开关，双端口各自通过开关连接到独立的计算机主机上。复杂可编程逻辑器件（Complex Programmable Logic Device，CPLD）或高级精简指令集机器（Advanced RISC Machines，ARM）作为独立的控制电路，确保双端口静态存储器的每一个端口上存在一个开关，两个开关不能同时闭合。当交换的内容是文件数据时，它确实给出了一种隔离断开

的实现。当交换的内容是 IP 包时，则不是。因为双端口 RAM 可以进行 IP 包的存储和转发，这是一种结构缺陷。

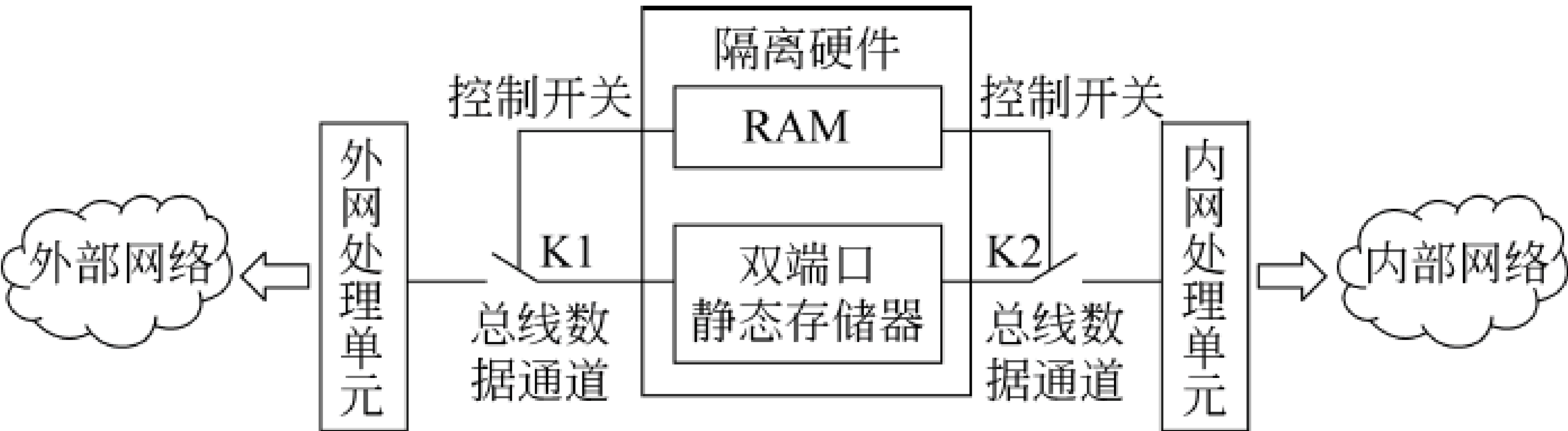



图 4.24　一种基于通信总线的网闸结构

采用这种技术的产品，应该严格检查是否实现了 TCP/IP 协议的剥离，是否实现了应用协议的剥离，确保是应用输出或输入的文件数据被转发，而不是 IP 包。除此之外，还必须有机制来保证双端口 RAM 不会被黑客用来转发 IP 包。如果设计不当，TCP/IP 协议没有剥离，IP 包会直接被写入内存存储介质，并且被转发。在这种情况下，尽管物理层是断开的，链路层也是断开的，由于 TCP/IP 协议的 3 层和 4 层没有断开，也不能算作网络隔离。

3）基于单向通道的网闸技术

由于双向通道可能会为攻击提供一种攻击通道，近年兴起了单向通道技术。为了说明这个问题，先分析一下下面的常见攻击过程。

（1）攻击者伪装攻击信息。

（2）伪装信息搭载正常数据通过网闸。

（3）伪装的攻击信息将自己还原，收集信息，并采用与（2）同样的手段向攻击者报告。

（4）攻击根据已经取得的权限进行下一步动作。

这样，双向的“摆渡”是无法切断这样的攻击通道的，即使采用了协议落地的手段也无济于事，因为某些攻击的行为还可能掩藏于“纯数据”之中，就像罪犯将毒品藏到体内一样。

如图 4.25 所示，单向通道就是把通信的收、发两个链路完全分开，在一个通道中不能完成通信的反馈，攻击行为就成了半开的连接，不能发挥效果。发送方只管发送数据，数据方只管接收数据。

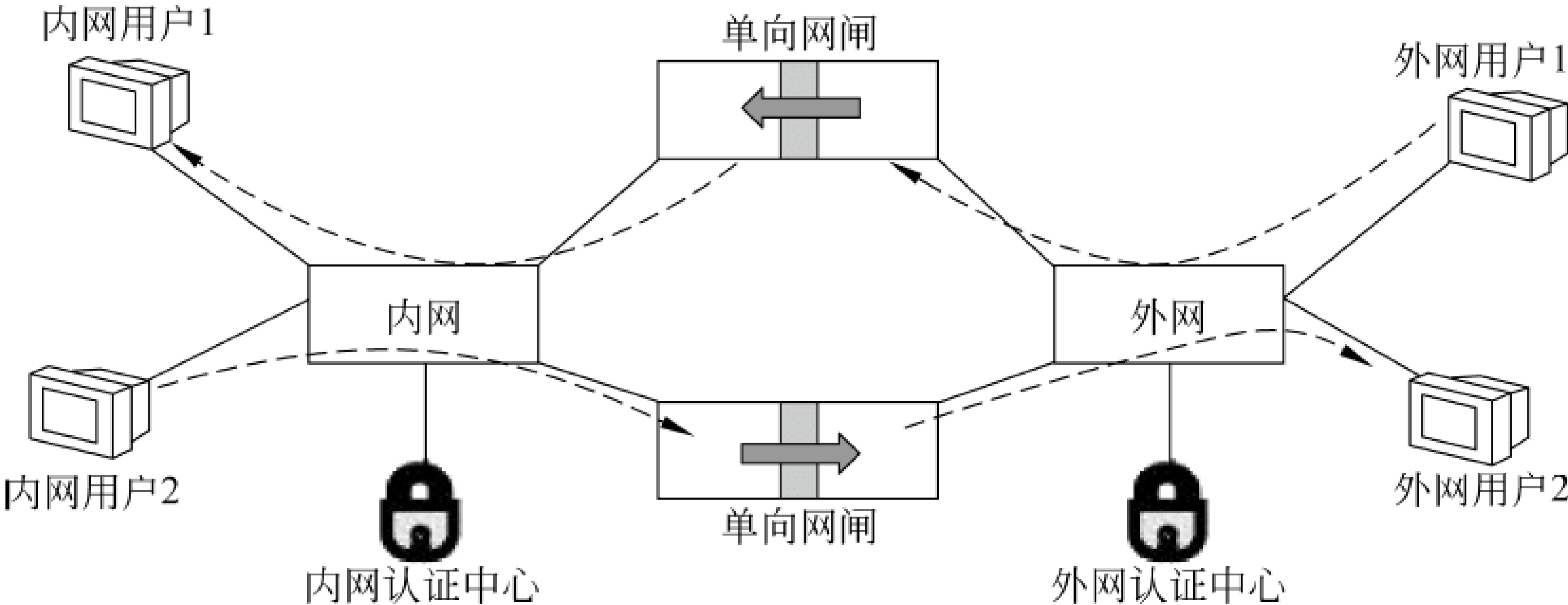



图 4.25　一种基于单向通道的网闸结构

单向通道技术没有差错重传机制，发送方并不知道接收方是否可靠地接收到数据，必须通过其他机制来提供可靠保障。例如，加上简单的控制信号，实现数据的差错重发等。

## 4.3 Internet 安全协议

当初，TCP/IP 开发的目标主要有两点：互连和高效，而没有考虑安全问题。随着 Internet 的广泛应用，安全问题逐步突出出来。为了解决 Internet 上的数据安全传输问题，人们采用了打“补丁”的办法：一块补丁打在 IP 层之上，称为 IPSec（IP Security）；一块补丁打在 TCP 层与应用层之间，称为安全套接层（Secure Socket Layer，SSL）。

它们采用了现代密码学方法，是在具体环境下实现机密性保护和认证性服务的范例。

### 4.3.1 IPSec

IP 是 TCP/IP 网络中最关键的一层，其安全性是整个 TCP/IP 安全的基础。为了弥补 IP 安全的缺陷，1994 年 IETF（Internet 安全任务组）成立了一个工作组来制定和推动关于 IP 安全的协议标准，并于 1995 年 5 月公布了一套 IETF 草案，形成一个 IP 安全体系。

#### 1．IP 安全分析

IP 层是关系整个 TCP/IP 安全的核心和基础。但是，由于当初设计时的环境和所考虑的基本出发点，对 IP 没有过多地考虑防卫问题，只是设法使网络能够方便地互连。这种不设防政策，给 Internet 造成许多安全隐患和漏洞，并随着攻击技术的提高，问题的严重性日益加剧。下面举几个例子来说明 IP 遭受的安全威胁。

（1）IPv4 缺乏对通信双方真实身份的验证能力，仅仅采用基于源 IP 地址的可认证机制，并且由于 IP 地址可以进行软件配置，这样就给攻击者以有机可乘，可以在一台计算机上假冒另一台计算机向接收方发送数据包，而接收方又无法判断接收到的数据包的真实性。这种 IP 欺骗可以在多种场合制造灾难。

（2）IPv4 缺乏对网络上传输的数据包进行机密性和完整性保护的能力，一般情况下 IP 包是明文传输的，第三方很容易窃听到 IP 数据包并提取其中的数据，甚至篡改窃取到的数据包内容，而且不被发觉，因为只要相应地修改校验和即可。

（3）由于数据包中没有携带时间戳、一次性随机数等，很容易遭受重放攻击。攻击者搜集特定 IP 包，进行一定处理就可以一一重新发送，欺骗对方。

（4）路由器布局是 Internet 的骨架。路由器不设防，将会使路由信息暴露，为攻击者提供入侵途径。

#### 2．IPSec 的传输模式和隧道模式

IPSec 是一套协议包，它提供了如图 4.26 所示的两种封装方式对 IP 数据包进行封装，形成了两种工作模式：传输模式（transport mode）和隧道模式（tunnel mode）。

1）传输模式

在传输模式中，IPSec 只保护 IP 数据包中的有效数据——传输层数据包，或者说只保护上层协议的数据包，即只对传输层数据包进行加密或认证，原来的 IP 头不变。所以 IPSec

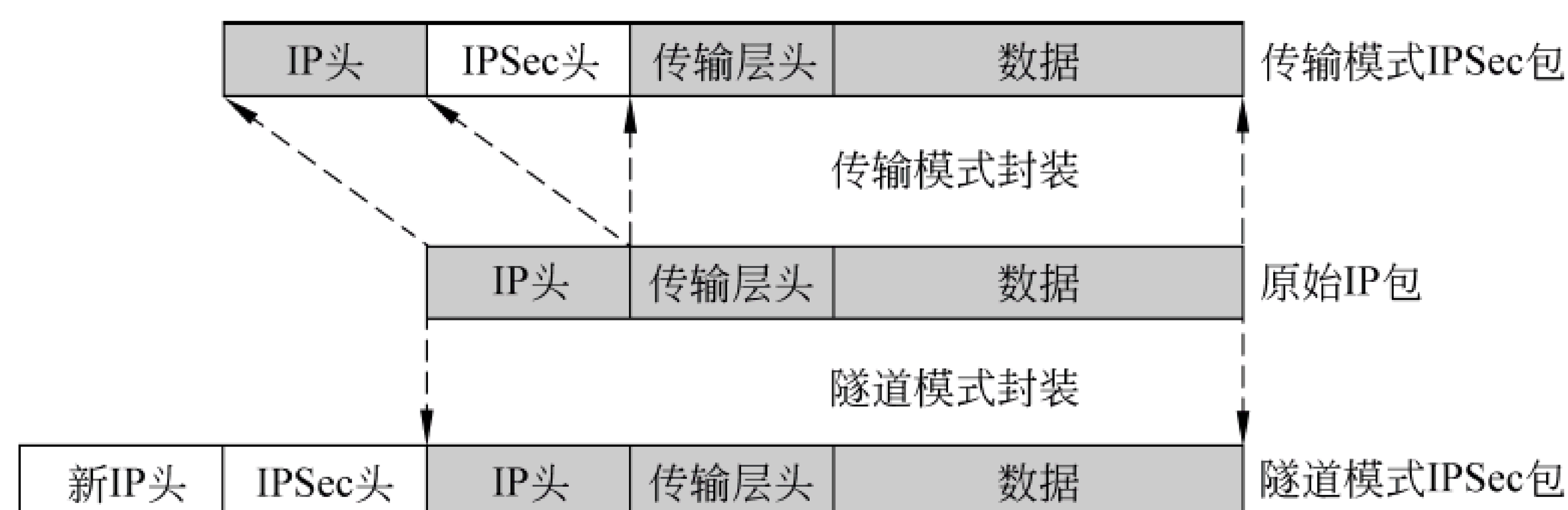



图 4.26 IPSec 两种封装的基本格式

头添加在 IP 头与 IP 数据之间。这样，封装了的数据包还可以直接传送到目的主机。这样就可以形成两台主机之间的安全通信。

传输模式的特点如下。

（1）只保护数据，不保护 IP 头，即 IP 地址是暴露的。

（2）用于两个主机之间。

2）隧道模式

隧道模式保护整个 IP 包，既保护 IP 包的内容，也保护 IP 包的头部，所以 IPSec 头添加在整个 IP 数据包之前。但这样，所有的路由器都不知道该数据包从哪里来，到哪里去了。为此还必须给新的受保护的数据包再加上一个新的 IP 头。在新的 IP 头中，不再使用原来的源地址和目的地址，只给出源端路由器地址和目的端路由器地址。也就是说，源 IP 数据包在源端路由器中被加密、被包装成新的 IP 数据包发送，在目的路由器中进行解包、解密，再从原来的 IP 头取出真实目的地址，将数据传送给目的主机。这样，监听者只能监听到两端路由器的地址，无法监听到数据包真实的源主机和目的主机的地址。这样就形成两台路由器之间的安全通信模式，这种数据包在传输的过程中，路由器只检查新的 IP 头。新的 IP 头定义了从源路由器到目的路由器之间的一条虚拟路径，好像在两个路由器之间形成一个可以安全通信的隧道。

总之，隧道模式有如下特点。

（1）既保护数据，又保护 IP 头。

（2）用于两个安全网关之间。

（3）无法控制来自内部的攻击。

### 3．AH 协议和 ESP 协议

IPSec 的核心是两个安全协议：认证头（Authentication Header，AH）协议和安全负荷封装（Encapsulating Security Payload，ESP）协议，它们分别提供了两种安全机制：认证和加密。它们都用来封装 IP 数据包。

1）AH 协议

AH 为 IP 包提供数据的完整性和验证服务。

（1）对数据使用完整性检查，可以判定数据包在传输过程中是否被修改。

（2）通过验证机制，终端系统或设备可以对用户或应用进行验证，并过滤通信流；还

可以防止地址欺骗和重放攻击。

AH 具有如图 4.27 所示的格式。




图 4.27　AH 格式

（1）下一个头（8b）：标识紧跟验证头的下一个头的类型。

（2）载荷长度（8b）：为以 32b 为单位的验证数据长度加 1。如默认的验证数据字段长度为 96b，为 3 个 32b；加上 1，得 4，即默认的验证数据的 AH 头的载荷长度为 4。

（3）保留（16b）：备以后使用。

（4）安全参数索引（32b）：用于标识一个安全协同（Security Association，SA）。

（5）序号（8b）：无符号单调递增计数值，用于 IP 数据包的重放检查。

（6）认证数据（32b 的整数倍可变长数据）：含数据包的 ICV（完整性校验值）或 MAC。

图 4.28 为在传输模式和隧道模式下的 AH 包格式。

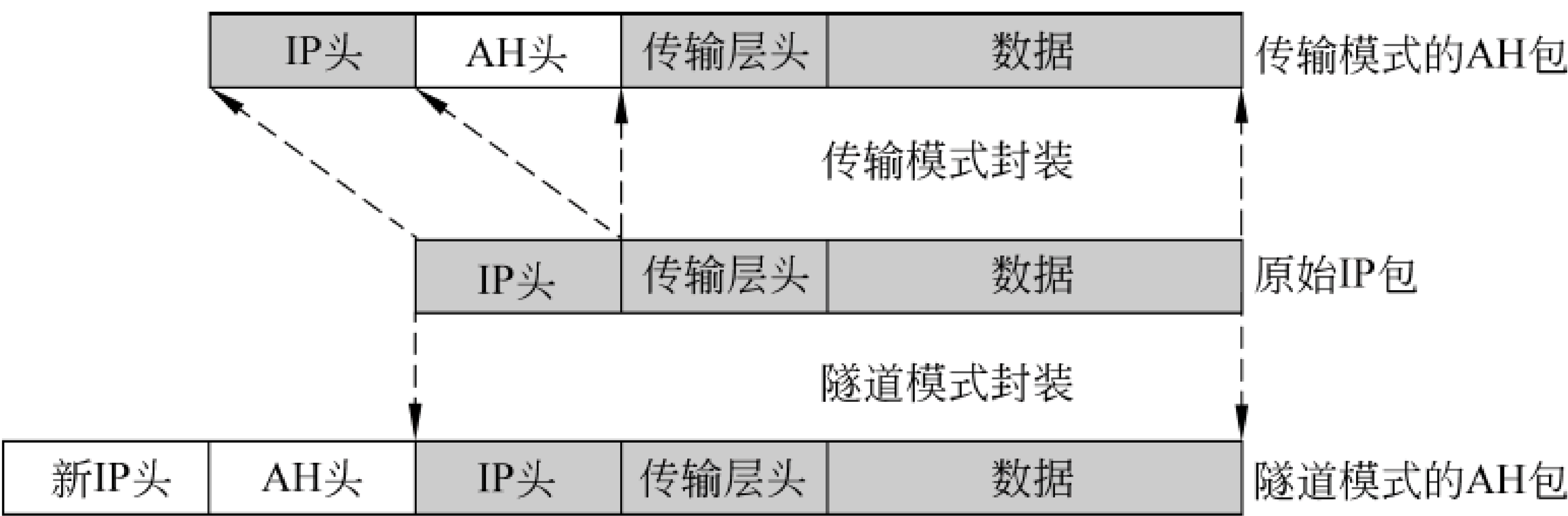



图 4.28　传输模式和隧道模式下的 AH 包格式

2）ESP 协议

ESP 协议为 IP 数据包提供如下服务。

（1）IP 包的机密性保护。

（2）数据源验证。

（3）数据完整性保护。

（4）抗重放保护。

ESP 具有如图 4.29 所示的格式。

其中：

① 下一个头（8b）：通过标识载荷中的第一个头（如 IPv6 中的扩展头，或诸如 TCP 等上层头）决定载荷数据字段中数据的类型。

② 安全参数索引（32b）：标识一个安全协同（SA）。

③ 序号（8b）：无符号单调递增计数值，用于 IP 数据包的重放检查。

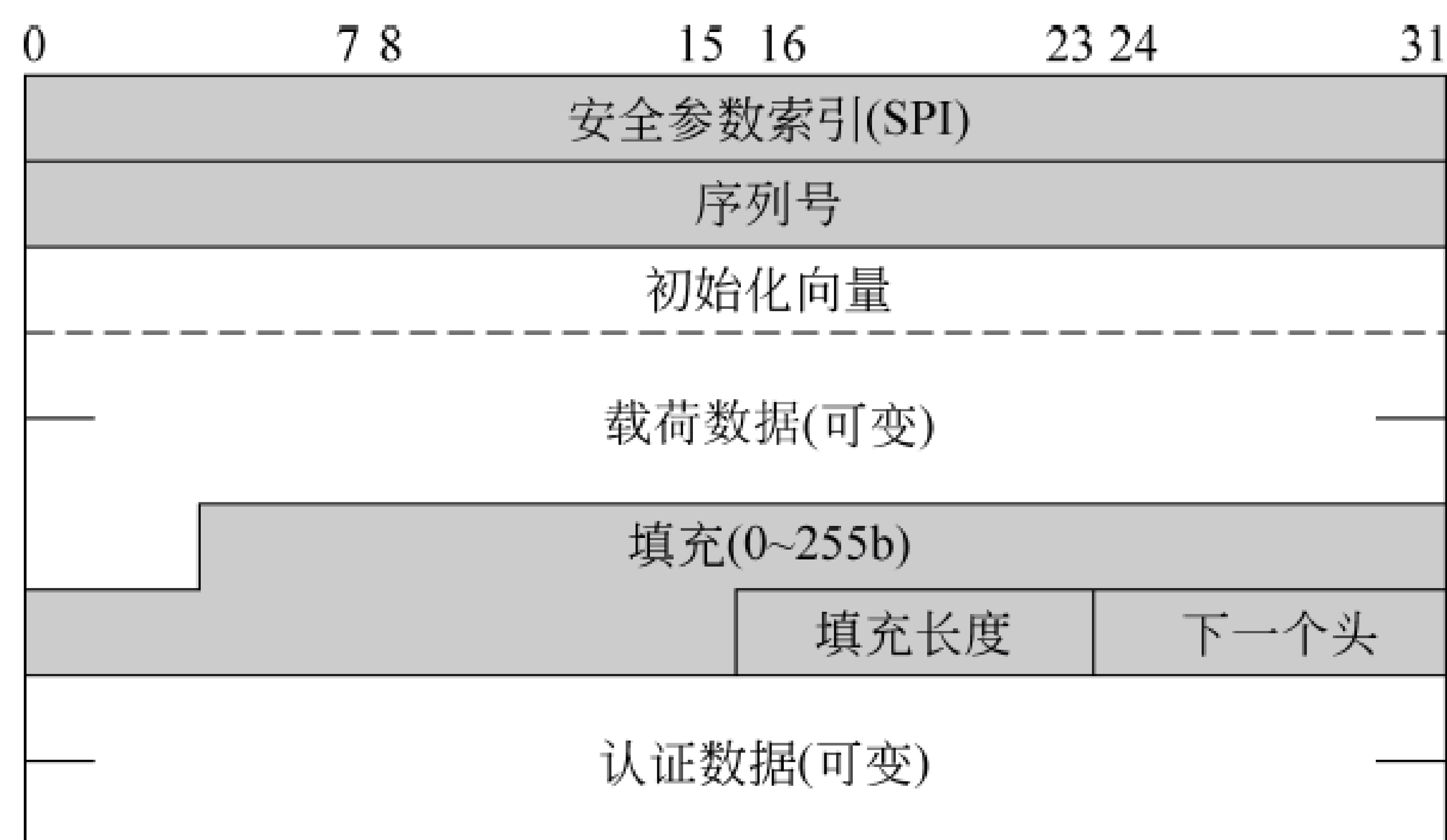



图 4.29　ESP 格式

④ 认证数据（32b 的整数倍的可变长数据）：用于填入 ICV（完整性校验值）。IVC 的计算范围为 ESP 包中除掉验证数据字段部分。

⑤ 填充项（0～255b）：额外字节。

⑥ 填充长度（8b）：填充的字节数。

⑦ 载荷数据（可变）：在传输模式下为传输层数据段，在隧道模式下为 IP 包。

图 4.30 为在传输模式和隧道模式下的 ESP 包格式。

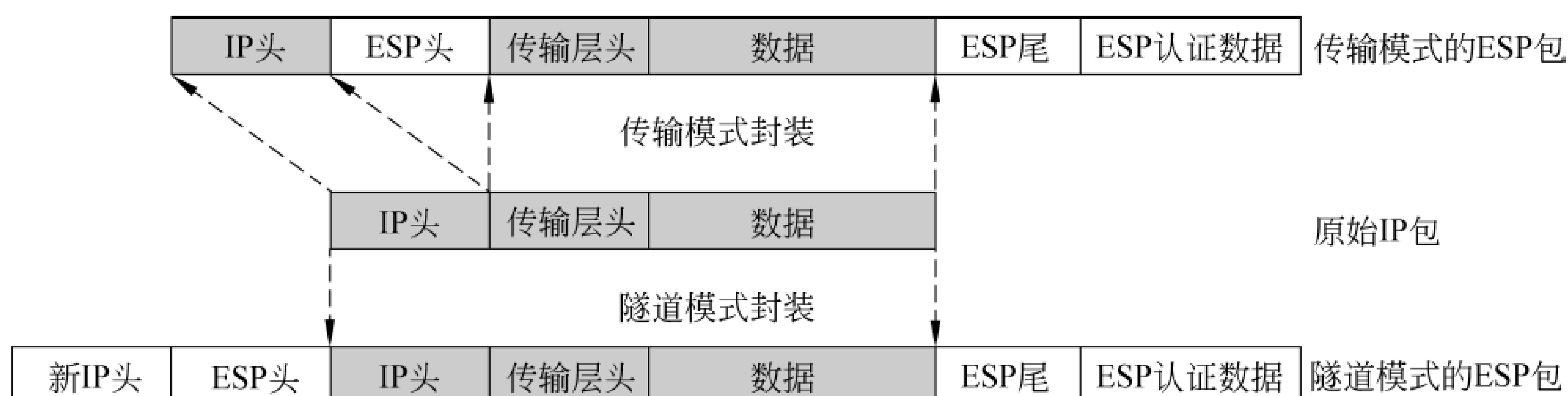



图 4.30　传输模式和隧道模式下的 ESP 包格式

### 4．IKE

IPSec 的密钥管理包括密钥的确定和分配，可以采用手工或自动方式进行。IPSec 默认的自动密钥管理协议是 Internet 密钥交换（Internet Key Exchange，IKE）。IKE 主要起两个作用。

（1）安全关联（Security Associations，SA）的集中化管理，以减少连接时间。

（2）密钥的生成与管理。

IKE 规定了验证 IPSec 对等实体、协商安全服务和生成会话密钥的方法。IKE 将密钥协商结果保留在 SA 中，供 AH 和 ESP 以后通信时使用。

IKE 有 4 种身份认证方式。

（1）基于数字签名的认证：利用数字证书表示身份，利用数字签名算法计算出一个签名来验证身份。

（2）基于公钥的认证：用对方的公钥加密身份，通过检查对方发来的哈希值进行认证。

（3）基于修正的公钥：对上一个方式的修正。

（4）基于预共享字符串：双方事先商定好一个共享的字符串。

### 5．IPSec 体系结构

IPSec 各部件之间的关系如图 4.31 所示。

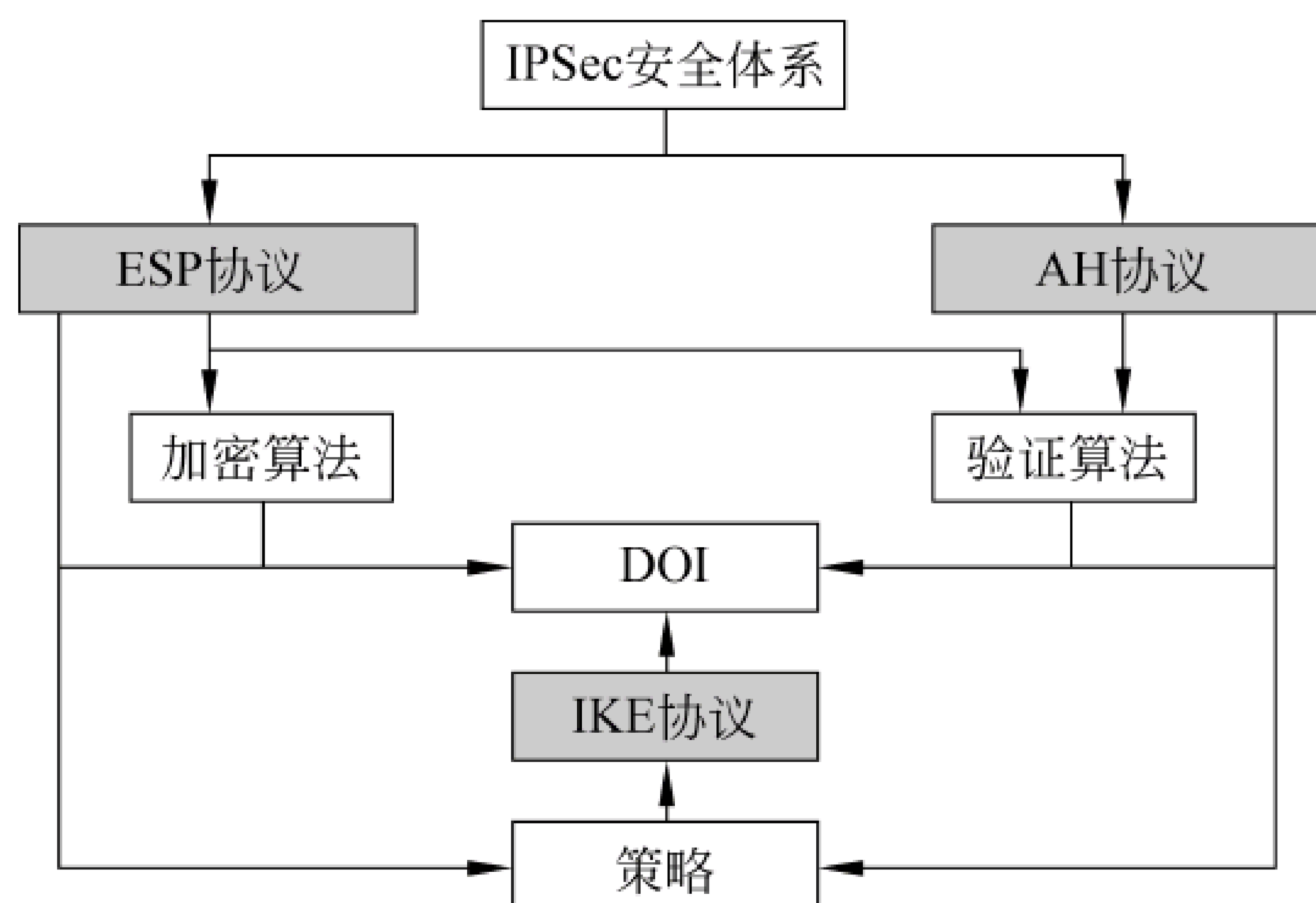



图 4.31　IPSec 各部件之间的关系

对 IPSec 的体系结构有以下几点说明。

（1）IPSec 的加密只用于 ESP。目前的 IPSec 标准要求任何 IPSec 实现都必须支持 3DES 和 AES（高级加密标准）。

（2）IPSec 的验证算法可用于 AH 和 ESP，主要采用 HMAC。HMAC 将消息和密钥作为输入来计算 MAC。MAC 保存在 AH/ESP 头中的验证数据字段中。目的地址收到 IP 包后，使用相同的验证算法和密钥计算一个新的 MAC，并与数据包中的 MAC 比对。

（3）解释域（Domain of Interpretation，DOI）用来组合相关协议，如定义负载的格式和交换的类型，以及对安全相关信息的命名约定，例如对安全策略或者加密算法和模式的命名等。

## 4.3.2　SSL

SSL 是 Netscape 公司提出的一种建构在 TCP 之上、为 Web 提供安全服务的安全标准。1999 年，SSL 被 IETF 接收后，经过改进以 TLS（Transport Layer Security）协议为名推出。于是，形成有专利保护的 SSL 和成为标准的 TLS 两种版本。不过，SSL 已经成为事实上的标准。虽然 SSL 的初衷是为 Web 提供安全服务，但是由于它与应用层协议无关性的开发思想，使其可以基于传输层为高层应用协议提供透明的安全服务。其中在 Web 服务器和浏览器之间的安全通信则是它最典型的应用，几乎所有 Web 服务器和浏览器都支持它，并且把基于 SSL 的 HTTP 协议称为 HTTPS 协议。

### 1．SSL 的工作过程

SSL 的设计目标是基于客户/服务器工作方式（点对点的信息传输），使高层协议在进行通信之前就能完成加密算法、密钥协商和服务器的认证，并在此基础上进行加密的安全通信（即对发送的消息数据进行分组、压缩、加密和生成认证码），为应用会话提供防窃听、防篡改和防消息伪造服务。所以它位于应用层和传输层之间。

实现上述目标的基本过程如下。

（1）安全协商：互相交换 SSL 版本号和所支持的加密算法等信息。

（2）彼此认证。

① 服务器将自己由 CA 的私钥加密的证书告诉浏览器。服务器也可以向浏览器发出证书请求，对浏览器进行认证。

② 浏览器检查服务器的证书（是否由自己列表中的某个 CA 颁发）：不合法，则终止连接；合法，则进入生成会话密钥步骤。

③ 如果服务器有证书请求，浏览器也要发送自己的证书。

（3）生成会话密钥。

① 浏览器用 CA 的公钥对服务器的证书解密，获得服务器的公钥。

② 浏览器生成一个随机会话密钥，用服务器的公钥加密后，发送给服务器。

（4）启动会话密钥。

① 浏览器向服务器发送消息，告诉服务器以后自己发送的信息将用协商好的会话密钥加密。

② 浏览器再向服务器发送一个加密消息，告诉服务器会话协商过程完成。

③ 服务器向浏览器发送消息，告诉浏览器以后自己发送的信息将用协商好的会话密钥加密。

④ 服务器再向浏览器发送一个加密消息，告诉浏览器会话协商过程完成。

（5）SSL 会话正式开始：双方用协商好的会话密钥加密发送的消息。

**2．SSL 体系结构**

为了实现上述过程，SSL 体系结构由两层组成。

1）握手层（管理层）

用于密钥的协商和管理，由握手协议、更改密码规范协议和警报协议组成。

（1）SSL 握手协议（handshake protocol）：准许服务器端与客户端在开始传输数据前可以通过特定的加密算法相互鉴别。

（2）SSL 更改密码规范协议（change cipher spec protocol）：保证可扩展性。

（3）SSL 警报协议（alert protocol）：产生必要的警报信息。

2）记录层

运行 SSL 记录协议（record protocol），为高层应用协议提供各种安全服务，对上层数据进行加密、产生 MAC 等并进行封装。

图 4.32 表明 SSL 在 TCP/IP 协议栈中的位置。它位于传输层之上、应用层之下，并独




图 4.32　SSL 体系结构

立于应用层，使应用层可以直接建立在 SSL 上。

**3．SSL 握手**

连接（connection）和会话（session）是 SSL 中的两个重要概念：一个 SSL 会话是客户机与服务器之间的一个关联，一个 SSL 连接提供一种合适的服务类型传输。SSL 连接是点对点的关系，并且连接是暂时的，每一个连接只与一个会话关联。会话定义了一组可供多个连接共享的加密安全参数，以避免为每一个连接提供新的安全参数所需的昂贵谈判代价。

客户机与服务器要建立一个会话，就必须握手。SSL 会话由 SSL 握手协议创建或恢复。图 4.33（a）为创建一个会话的握手过程，图 4.33（b）为恢复一个会话的握手过程。

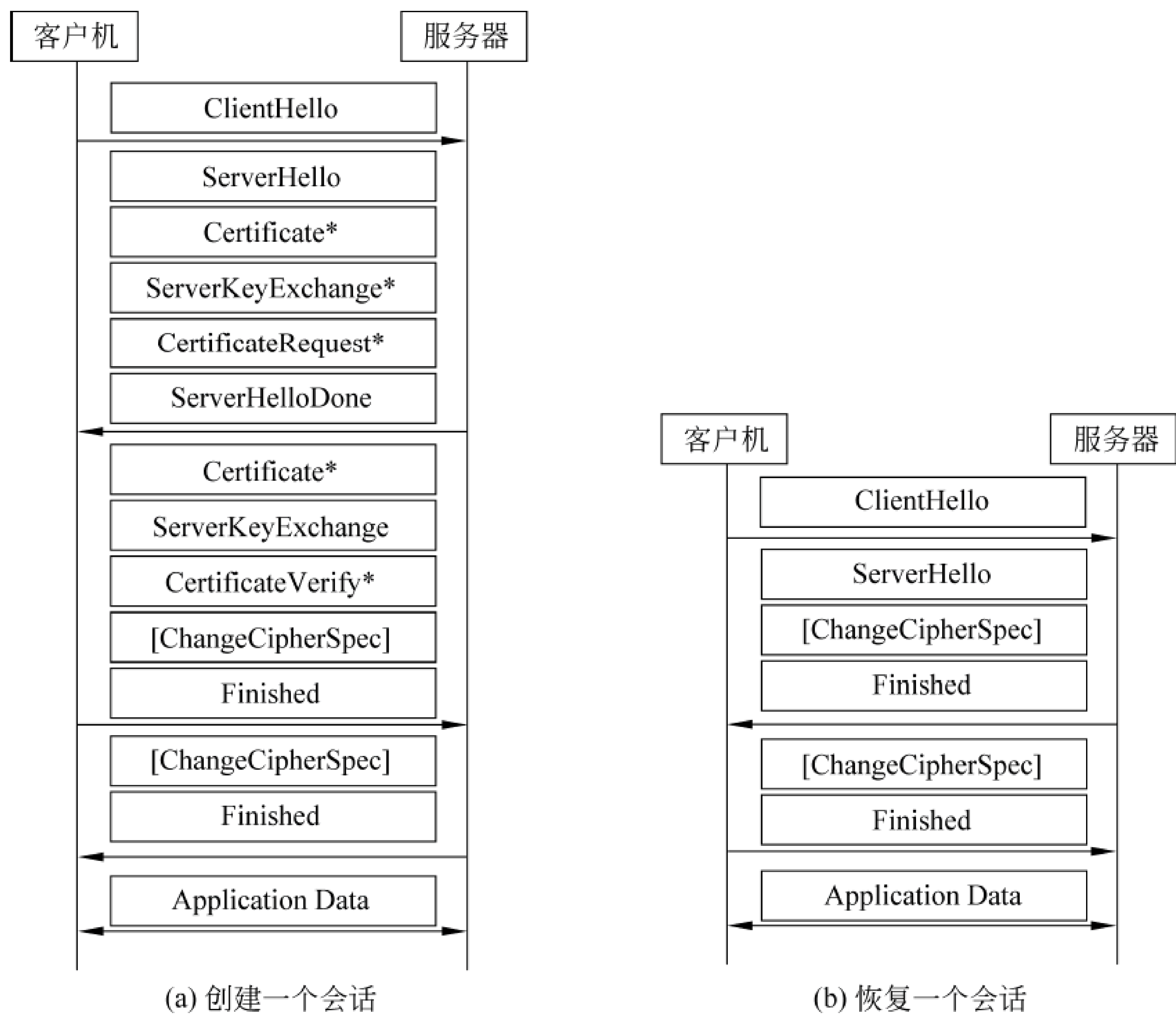



图 4.33　SSL 握手过程

注：* 表示依当时情形，可选择性发出。

下面主要介绍创建会话时的握手过程。

1）Hello 阶段

握手协议从客户机发出的第一道信息 ClientHello 开始。

（1）ClientHello 和 ServerHello，用于协商安全参数，包括协议版本号、会话识别码（session_id）、时间戳、密码算法协商（cipher suit）、压缩算法、两个 28B 的随机数（ClientHello.random 和 ServerHello.random）。

（2）Certificate，密钥交换信息，在要验证服务器时发出。

（3）ServerKeyExchange，送出供客户机计算出共享密钥的参数，包含服务器临时公钥。这些信息一般包含在 Certificate 中。只在下列情况下才由服务器发出。

① 不需要验证服务器。

② 要求验证服务器，但服务器无证书或服务器证书是用于签名。

（4）CertificateRequest 是在服务器要求验证客户机时发出。

（5）ServerHelloDone 表示双方握手过程的 Hello 阶段结束。

这时，服务器等待客户机回音。

2）加解密参数传输

（1）Certificate，回答服务器的 CertificateRequest 要求的信息。服务器无要求时，不发。

（2）ClientKeyExchange，对 ClientHello 和 ServerHello 密钥交换算法的回复，以 ServerKeyExchange 所选的算法进行，让双方可以共享密钥。

（3）CertificateVerify，对此前服务器送来的所有信息（ClientHello/ServerHello、Certificate 和 ServerKeyExchange）产生的签名，让服务器进一步确定客户机的正确性。

（4）ExchangeCipherSpec，SSL 更改密码规范协议消息。

（5）Finished，用协商好的算法和密钥加密的握手完成消息。

3）服务器确认

（1）ExchangeCipherSpec，回复客户机的 ExchangeCipherSpec 消息。

（2）Finished，用协商好的算法和密钥加密的握手完成消息。

4）开始传输应用数据

恢复一个已经存在的会话时，握手过程一般只需要 Hello 阶段。

**4．SSL 记录协议的封装**

在 SSL 体系中，当上层（应用层或表示层）的应用要选用 SSL 协议时，上层（握手、警报、更改密码规范、HTTP 等）协议信息会通过 SSL 记录子协议使用一些必要的程序将加密码、压缩码和 MAC 等封装成若干数据包，再通过其下层（基本上都是从呼叫 socket 接口层）传送出去。

记录协议的封装过程如图 4.34 所示。

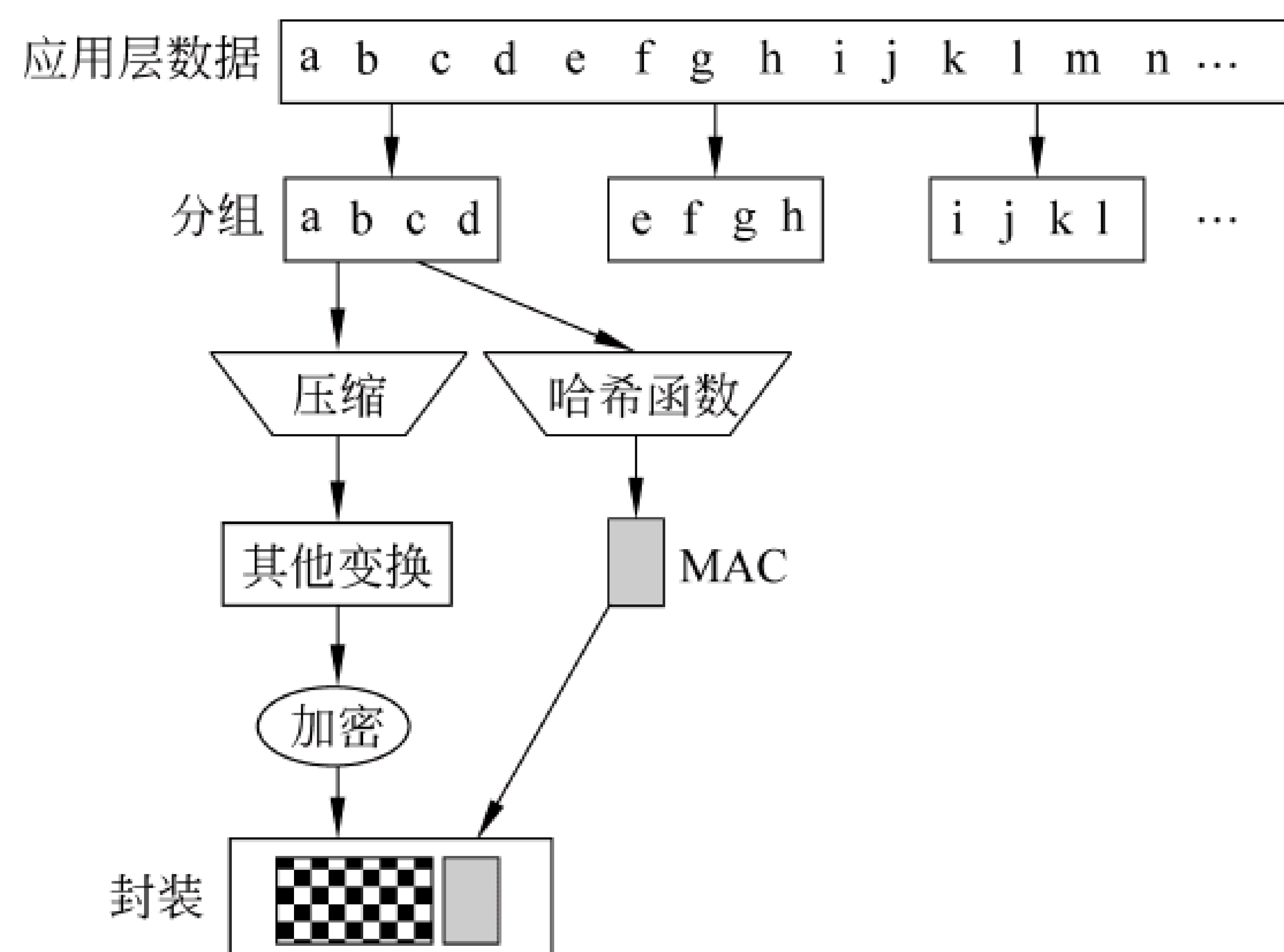



图 4.34　记录协议的封装过程

### 4.3.3 VPN

#### 1．VPN 的基本原理

虚拟专用网（Virtual Private Network，VPN）是指将物理上分布在不同地点的专用网络通过不可信任的公共网络构造成逻辑上信任的虚拟子网，进行安全的通信。这里公共网络主要指 Internet。

图 4.35 为 VPN 的结构示意图。图中有 3 个专网，它们都位于 VPN 设备的后面，同时由路由器连接到公共网。VPN 技术采用了安全封装、加密、认证、存取控制、数据完整性保护等措施，使得敏感信息只有预定的接收者才能读懂，实现信息的安全传输，使信息不被泄露、篡改和复制，相当于在各 VPN 设备间形成一些跨越 Internet 的虚拟通道——“隧道”。

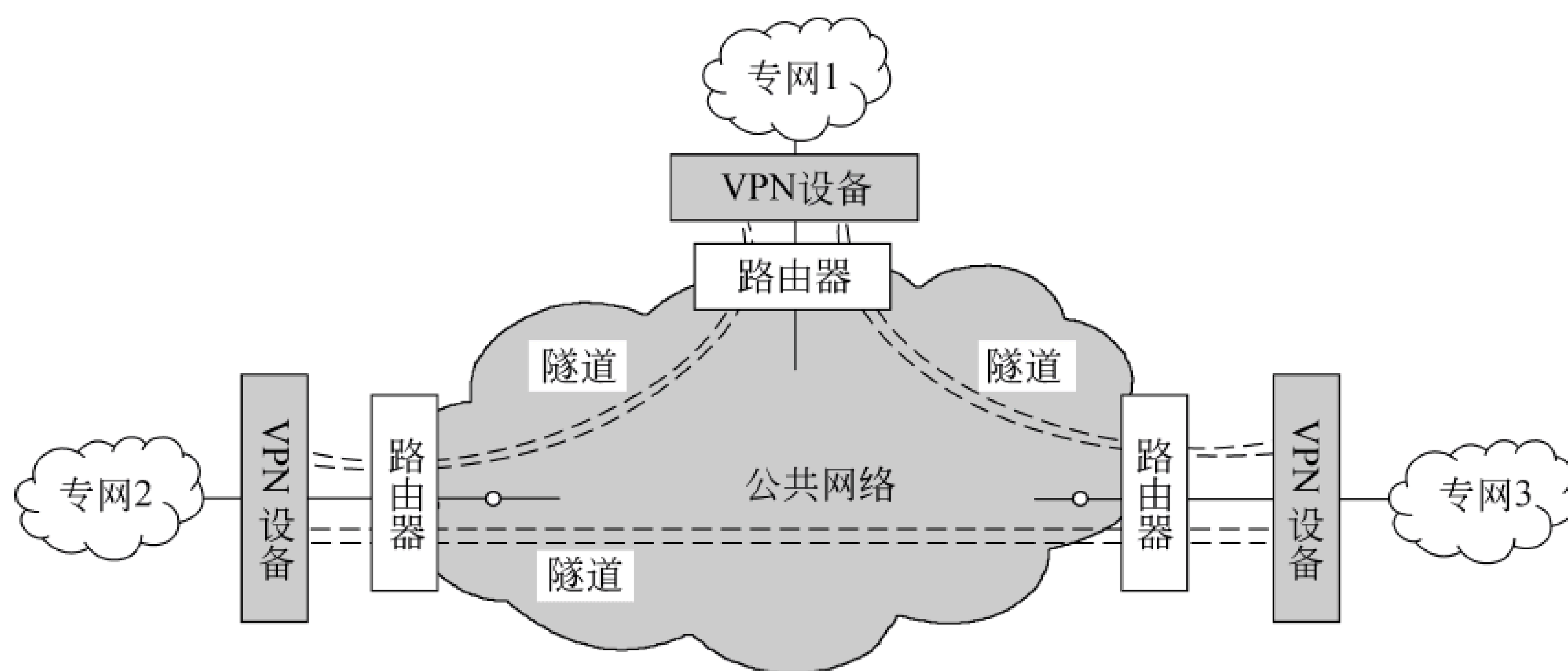



图 4.35　VPN 的结构与基本原理

隧道的建立主要有两种方式：客户启动（client-initiated）和客户透明（client-transparent）。客户启动也称为自愿型隧道，要求客户和服务器（或网关）都安装特殊的隧道软件，以便在 Internet 中可以任意使用隧道技术，完全由自己控制数据的安全。客户透明也称为强制型隧道，只需要服务器端安装特殊的隧道软件，客户软件只用来初始化隧道，并使用用户 ID、口令或数字证书进行权限鉴别，使用起来比较方便，主要供 ISP 将用户连接到 Internet 时使用。

VPN 的基本处理过程如下。

（1）要保护的主机发送明文信息到其 VPN 设备。

（2）VPN 设备根据网络管理员设置的规则，确定是对数据进行加密还是直接传送。

（3）对需要加密的数据，VPN 设备将其整个数据包（包括要传送的数据、源 IP 地址和目的 IP 地址）进行加密并附上数字签名，加上新的数据报头（包括目的 VPN 设备需要的安全信息和一些初始化参数），重新封装。

（4）将封装后的数据包通过隧道在公共网上传送。

（5）数据包到达目的 VPN 设备，将数据包解封，核对数字签名无误后，对数据包解密。

#### 2. 隧道结构

隧道技术是 VPN 技术的核心，它涉及数据的封装，可以利用 TCP/IP 协议作为主要传送协议，以一种安全的方式在公用网络（如 Internet）上传送。

在 VPN 中，双方的通信量很大，并且往往很熟悉，这样就可以使用复杂的专用加密和认证技术对通信双方的 VPN 进行加密和认证。为了实现这些功能，隧道被构造为一种 3 层结构。

（1）最底层是传输。传输协议用来传输上层的封装协议，IP、ATM、PVC 和 SVC 都是非常合适的传输技术。其中因为 IP 具有强大的路由选择能力，可以运行于不同的介质上，因而应用最为广泛。

（2）第二层是封装。封装协议用来建立、保持和拆卸隧道，或者说是数据的封装、打包与拆包。

（3）第三层是认证。

### 3．VPN 实现技术

目前，实现 VPN 的主要技术有两种：一种是基于 IPSec 协议的 VPN 模式；另一种是基于 SSL 协议的 VPN 模式。关于它们的具体实现方法，这里不再赘述。

### 4．VPN 的服务类型

从应用的角度看，VPN 的服务大致有如下 3 种类型。

（1）远程访问 VPN（access VPN）：适合在外地有流动办公的情况。这时的驻外工作人员只能通过宾馆或其他的设施以拨号方式与本部进行 VPN 连接，利用 HTTP、FTP 等或其他网络服务与本部交换信息。

（2）内联 VPN（intranet VPN）：适合在外地有固定分支机构的情形。这时，驻外分支机构通过 ISP 与本部进行 VPN 安全连接。

（3）外联 VPN（extranet VPN）：适合与业务伙伴之间通信的连接。这时，往往需要通过专线连接公共基础设施，并借助电子商务软件等与本部进行 VPN 连接。

## 实验 10　实现一个 VPN 连接

### 1. 实验目的

（1）理解 VPN 的工作原理。

（2）了解 VPN 的应用。

（3）掌握 VPN 实现的技术方法。

### 2. 实验内容

（1）实现一个 VPN 连接。

（2）测试连接后的 VPN 网络。

### 3. 建议环境

（1）在 Windows 操作系统中利用 PPTP 配置 VPN 网络，即在 Windows 2000 Server 中选择“开始”→“程序”→“管理工具”，单击“路由和远程访问”。

（2）在 Windows 中配置 IPSec，即选择“开始”→“程序”→“管理工具”，进入“本地安全策略”界面。在右侧窗口中，可以看到默认情况下 Windows 内置的“安全服务器”“客户端”和“服务器”3 个安全选项，并附有描述。

（3）在 Linux 操作系统中利用 CIPE 配置 VPN。CIPE（Crypto IP Encapsulation）是主要为 Linux 而开发的 VPN 实现软件。CIPE 使用默认的 CIPE 加密机制（标准的 Blowfish 或 IDEA 加密算法）来加密 IP 分组，并为这些分组添加目标头信息后，封装或“包围”在数据报（UDP）中。然后，这些 UDP 分组再通过 CIPE 虚拟网络设备（cipcbX）和 IP 层。

#### 4. 实验准备

（1）对要使用的 VPN 连接进行需求分析。

（2）根据需求分析提出使用的 VPN 连接方案。

（3）设计实现确定的 VPN 方案所需要的软硬件环境。

（4）设计进行 VPN 连接的步骤。

（5）设计进行 VPN 连接测试的方法和步骤。

#### 5. 推荐的分析讨论内容

（1）搜集各种 VPN 实现技术，并进行比较。

（2）对自己实现的 VPN 进行安全风险分析，提出改进设想。

（3）有的计算机网络具有单入口点，即出入网络的所有数据都只通过单个网关（路由器/防火墙），而有的网络中使用了多个网关。在这两种情况下，进行 VPN 配置有什么区别？

（4）其他发现或想到的问题。

## 4.4　入侵检测系统

### 4.4.1　入侵检测及其模型

#### 1. 入侵检测与入侵检测系统

入侵检测系统（Intrusion Detection System，IDS）是对计算机和网络系统资源上的恶意行为进行识别和响应的处理系统；它像雷达警戒一样，在不影响网络性能的前提下，对网络进行警戒、监控，从计算机网络的若干关键点收集信息，通过分析这些信息，看看网络中是否有违反安全策略的行为和遭到攻击的迹象，从而扩展了系统管理员的安全管理能力，提高了信息安全基础结构的完整性。

IDS 最早于 1980 年 4 月由 James P. Anderson 在为美国空军起草的技术报告 *Computer Security Threat Monitoring and Surveillance*（《计算机安全威胁监控与监视》）中提出。报告中提出了一种对计算机系统风险和威胁的分类方法，将威胁分为外部渗透、内部渗透和不法行为；提出了利用审计跟踪数据监视入侵活动的思想。这份报告被认为是入侵检测的开山之作。

这里，“入侵”（intrusion）是一个广义的概念，不仅包括发起攻击的人（包括黑客）取

得超出合法权限的行为，也包括收集漏洞信息，造成拒绝访问（DoS）等对系统造成危害的行为。而入侵检测（intrusion detection）就是对入侵行为的发觉。它通过对计算机网络等信息系统中若干关键点的有关信息的收集和分析，从中发现系统中是否存在违反安全规则的行为和被攻击的迹象。入侵检测系统就是进行入侵检测的软件和硬件的组合。

入侵检测作为一种积极主动的安全防护技术，提供了对内部攻击、外部攻击和误操作的实时保护，被认为是防火墙后面的第二道安全防线。

#### 2. 入侵检测系统的功能

具体说来，入侵检测系统的主要功能如下。

（1）监视并分析用户和系统的行为。

（2）审计系统配置和漏洞。

（3）评估敏感系统和数据的完整性。

（4）识别攻击行为，对异常行为进行统计。

（5）自动收集与系统相关的补丁。

（6）审计、识别并跟踪违反安全法规的行为。

（7）使用诱骗服务器记录黑客行为。

⋮

#### 3. 实时入侵检测和事后入侵检测

实时入侵检测在网络的连接过程中进行，通过攻击识别模块对用户当前的操作进行分析，一旦发现攻击迹象就转入攻击处理模块，如立即断开攻击者与主机的连接、收集证据或实施数据恢复等。如图 4.36 所示，这个检测过程是反复循环进行的。

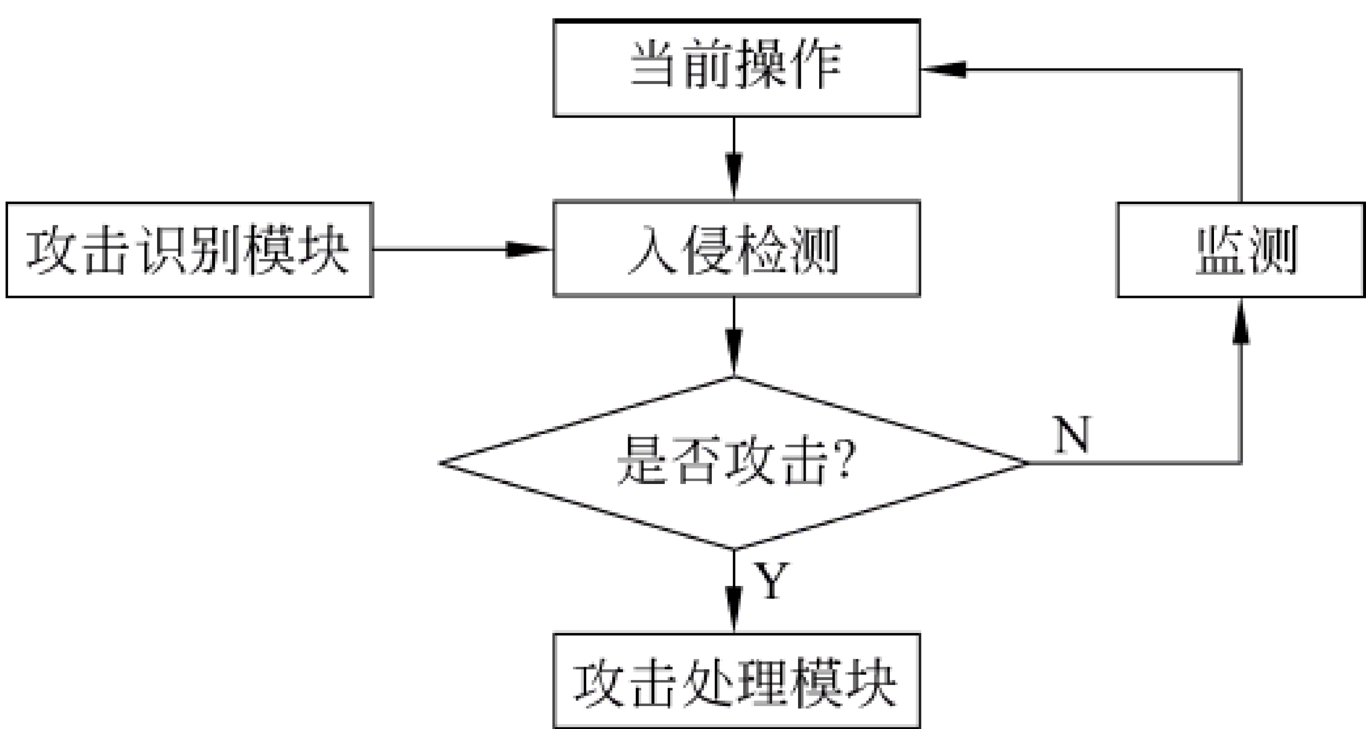



图 4.36　实时入侵检测过程

事后入侵检测是根据计算机系统对用户操作所做的历史审计记录，判断是否发生了攻击行为，如果有，则转入攻击处理模块处理。事后入侵检测通常由网络管理人员定期或不定期地进行。图 4.37 为事后入侵检测的过程。

#### 4. 入侵检测系统的基本结构

入侵检测是防火墙的合理补充，帮助系统对付来自外部或内部的攻击，扩展了系统管理员的安全管理能力（如安全审计、监视、攻击识别及其响应），提高了信息安全基础结构

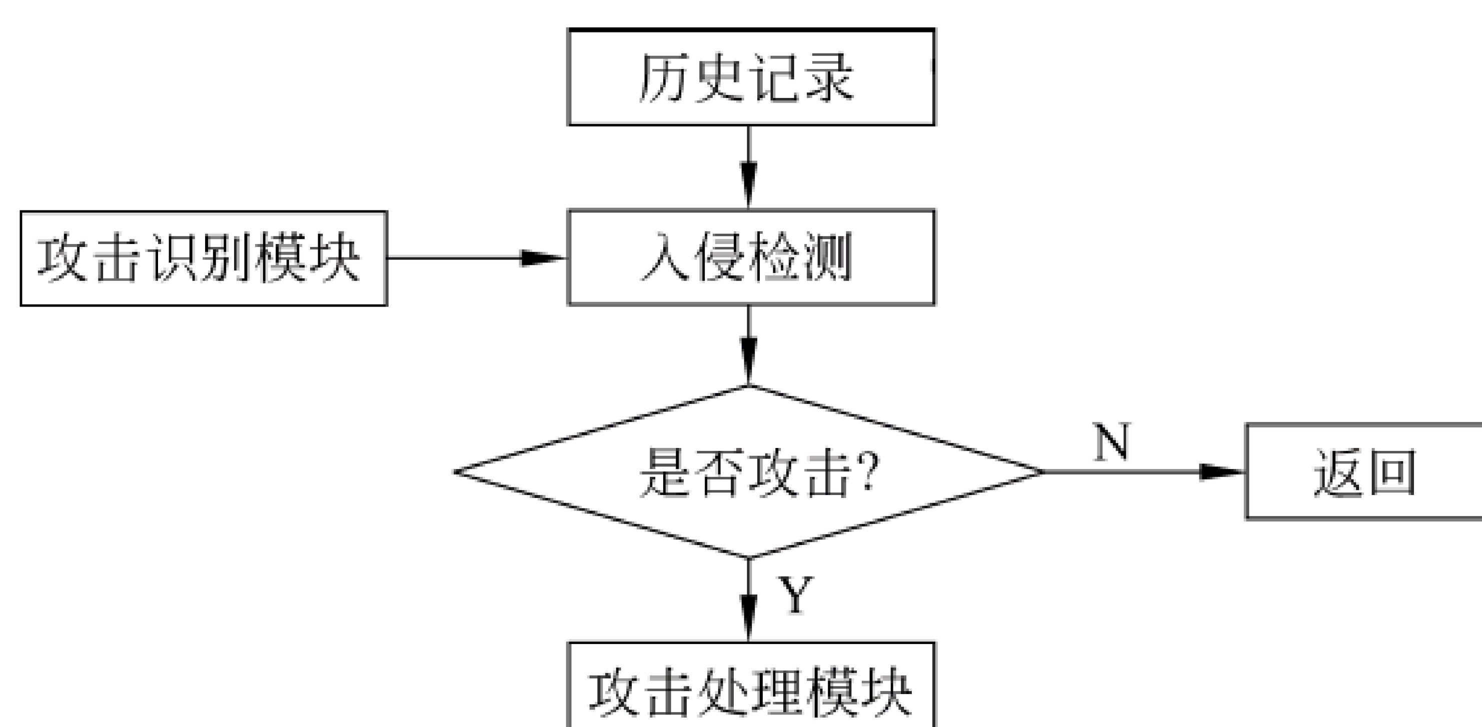



图 4.37　事后入侵检测的过程

的完整性。如图 4.38 所示，入侵检测系统的主要工作就是从信息系统的若干关键点上收集信息，然后分析这些信息，用来得到网络中有无违反安全策略的行为和遭到袭击的迹象。

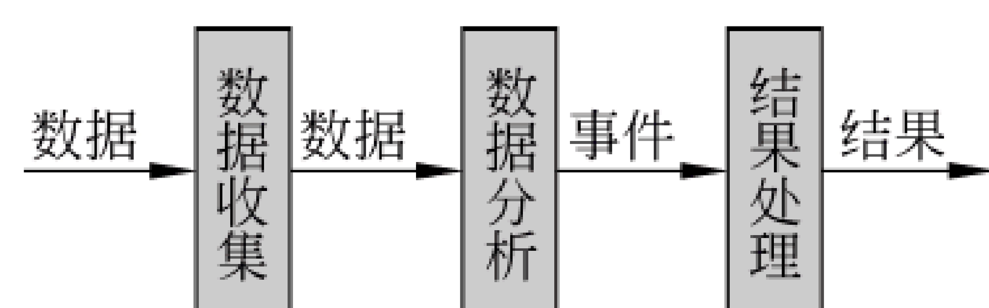



图 4.38　入侵检测系统的通用模型

入侵检测系统的这个模型比较粗略。但是它表明数据收集、数据分析和处理响应是一个入侵检测系统的最基本部件。

## 4.4.2　信息收集与数据分析

入侵检测的第一步是在信息系统的一些关键点上收集信息。这些信息就是入侵检测系统的输入数据。

### 1. 数据收集的内容

入侵检测系统收集的数据一般有如下 4 个方面。

1）主机和网络日志文件

主机和网络日志文件中记录了各种行为类型，每种行为类型又包含不同的信息。例如，记录“用户活动”类型的日志，就包含登录、用户 ID 改变、用户对文件的访问、授权和认证信息等内容。这些信息包含了发生在主机和网络上的不寻常和不期望活动的证据，留下黑客的踪迹。通过查看日志文件，能够发现成功的入侵或入侵企图，并很快地启动响应的应急响应程序。因此，充分利用主机和网络日志文件信息是检测入侵的必要条件。

2）目录和文件中的不期望的改变

网络环境中的文件系统包含很多软件和数据文件。包含重要信息的文件和私密数据文件经常是黑客修改或破坏的目标。黑客经常替换、修改和破坏他们获得访问权的系统上的文件，同时为了隐蔽他们在系统中的活动痕迹，还会尽力替换系统程序或修改系统日志文件。因此，目录和文件中的不期望的改变（包括修改、创建和删除），特别是那些正常情况下限制访问的对象，往往就是入侵发生的指示和信号。

3）程序执行中的不期望行为

每个在系统上执行的程序由一到多个进程来实现。每个进程都运行在特定权限的环境中，进程的行为由它运行时执行的操作来表现，这种环境控制着进程可访问的系统资源、程序和数据文件等；操作执行的方式不同，利用的系统资源也就不同。

操作包括计算、文件传输、设备与网络间其他进程的通信。黑客可能会将程序或服务的运行分解，从而导致它的失败，或者是以非用户或管理员意图的方式操作。因此，一个进程出现了不期望的行为可能表明黑客正在入侵本系统。

4）物理形式的入侵信息

黑客总是想方设法（如通过网络上由用户私自加上去的不安全（即未授权的）设备）去突破网络的周边防卫，以便能够在物理上访问内部网，在内部网上安装他们自己的设备和软件。例如，用户在家里可能安装调制解调器以访问远程办公室，那么这一拨号访问就成了威胁网络安全的后门。黑客就会利用这个后门来访问内部网，从而越过了内部网络原有的防护措施，然后捕获网络流量，进而攻击其他系统，并偷取敏感的私有信息等。

**2. 入侵检测系统的数据收集机制**

准确性、可靠性和效率是入侵检测系统数据收集机制的基本指标，在 IDS 中占据着举足轻重的位置。如果收集的数据时延较大，检测就会失去作用；如果数据不完整，系统的检测能力就会下降；如果由于错误或入侵者的行为致使收集的数据不正确，IDS 就会无法检测到某些入侵，给用户以安全的假象。

1）基于主机的数据收集和基于网络的数据收集

基于主机的 IDS 是在每台要保护的主机后台运行一个代理程序，检测主机运行日志中记录的未经授权的可疑行径，检测正在运行的进程是否合法并及时做出响应。

基于网络的入侵检测系统是在连接过程中监视特定网段的数据流，查找每一数据包内隐藏的恶意入侵，对发现的入侵做出及时的响应。在这种系统中，使用网络引擎执行监控任务。图 4.39 中给出了网络引擎所处的几个可能位置。

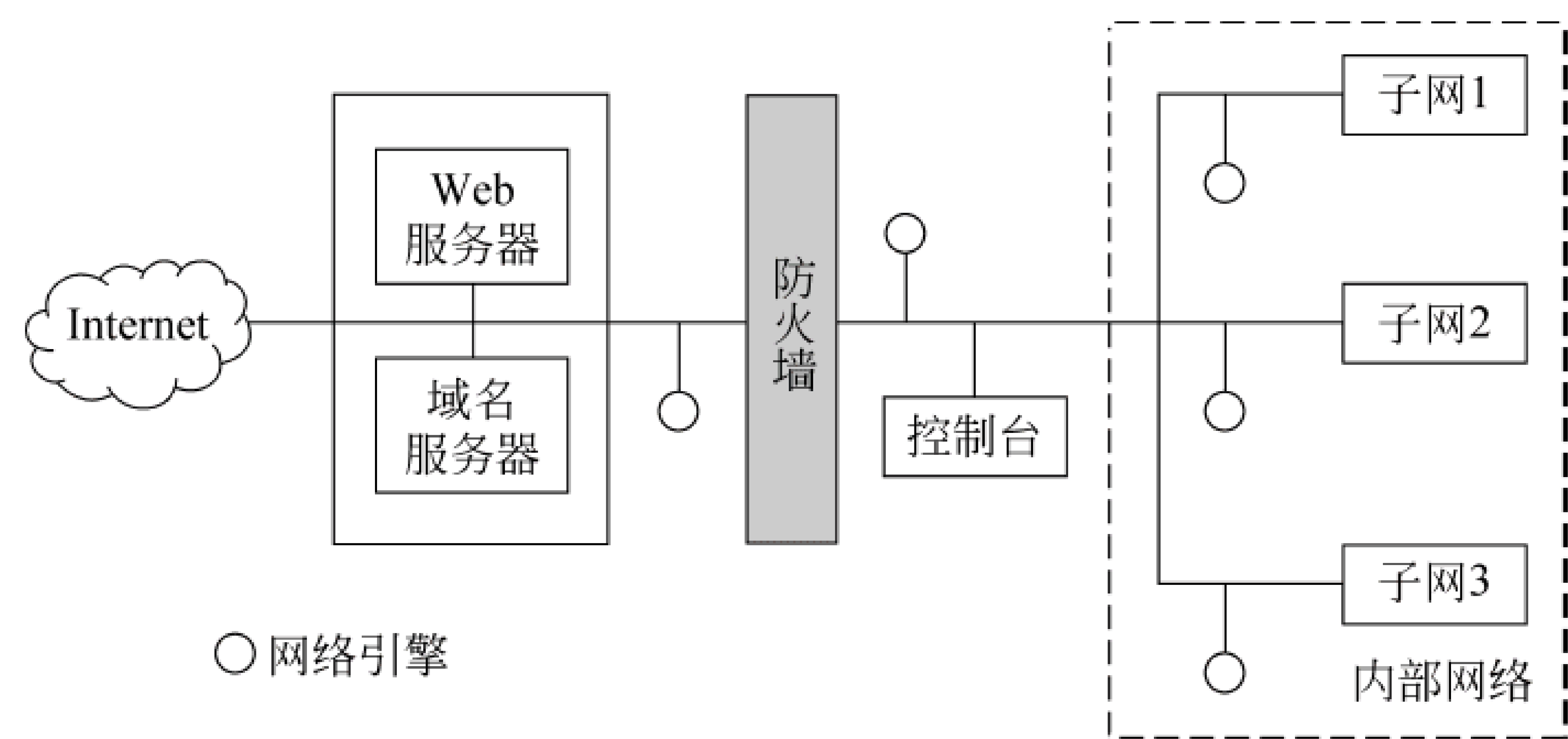



图 4.39　基于网络的 IDS 中网络引擎的配置

网络引擎所处位置不同，所起的作用不同。

（1）网络引擎配置在防火墙内，可以监测渗透过防火墙的攻击。

（2）网络引擎配置在防火墙外的非军事区，可以监测对防火墙的攻击。

（3）网络引擎配置在内部网络的各临界网段，可以监测内部的攻击。

控制台用于监控全网络的网络引擎。为了防止假扮控制台入侵或拦截数据，在控制台与网络引擎之间应创建安全通道。

基于网络的入侵检测系统主要用于实时监控网络关键路径。它的隐蔽性好、视野宽、侦测速度快、占用资源少、实施简便，并且还可以用单独的计算机实现，不增加主机负担。但难于发现所有数据包，对于加密环境无能为力，用在交换式以太网上比较困难。

基于主机的 IDS 提供了基于网络的 IDS 不能提供的一些功能，如二进制完整性检查、记录分析和非法进程关闭等。同时由于不受交换机隔离的影响，在交换网络中非常有用。但是它对网络流量不敏感，并且由于运行在后台，不能访问被保护系统的核心功能（不能将攻击阻挡在协议层之外）。它的内在结构没有任何束缚，并可以利用操作系统提供的功能，结合异常分析，较准确地报告攻击行为，而不是根据网上收集到的数据包去猜测发生的事件。但是它们往往要求为不同的平台开发不同的程序，从而增加了主机的负担。

总的看来，单纯地使用基于主机的入侵检测或基于网络的入侵检测，都会造成主动防御体系的不全面。但是，由于它们具有互补性，所以将两种产品结合起来，无缝地部署在网络内，就会构成综合了两者优势的主动防御体系，既可以发现网段中的攻击信息，又可以从系统日志中发现异常情况。这种系统一般为分布式，由多个部件组成。

2）分布式与集中式数据收集机制

分布式 IDS 收集的数据来自一些固定位置，而与受监视的网元数量无关。集中式 IDS 收集的数据来自一些与受监视的网元数量有一定比例关系的位置。

3）直接监控和间接监控

IDS 从它所监控的对象处直接获得数据，称为直接监控；反之，如果 IDS 依赖一个单独的进程或工具获得数据，则称为间接监控。

就检测入侵行为而言，直接监控要优于间接监控，这是因为：

（1）从非直接数据源获取的数据在被 IDS 使用之前，入侵者还有进行修改的潜在机会。

（2）非直接数据源可能无法记录某些事件，例如，它无法访问监视对象的内部信息。

（3）在间接监控中，数据一般都是通过某种机制（如编写审计代码）生成的，但这些机制并不满足 IDS 的具体要求，因而从间接数据源获得的数据量要比从直接数据源所获得的大得多。并且间接监控机制的可伸缩性小，一旦主机及其内部被监控要素增加，过滤数据的开销就会降低监控主机的性能。

（4）间接数据源的数据从产生到 IDS 访问之间有一个时延。

但是由于直接监控操作的复杂性，目前的 IDS 产品中只有不足 20%使用了直接监控机制。

4）外部探测器和内部探测器

外部探测器的监控组件（程序）独立于被监测组件（硬件或软件）。内部探测器的监控组件（程序）附加于被监测组件（硬件或软件）。表 4.10 给出了它们的优缺点比较。

### 3. 数据分析

数据分析是 IDS 的核心，其功能就是对从数据源提供的系统运行状态和活动记录进行同步、整理、组织、分类以及各种类型的细致分析，提取其中包含的系统活动特征或模式，用于对正常和异常行为的判断。

表 4.10　外部探测器和内部探测器的优缺点

| 比较内容 | 外部探测器 | 内部探测器 |
| --- | --- | --- |
| 错误引入和安全性 | • 代理消耗了过量资源<br>• 库调用错误地修改了某些参数<br>• 有被入侵者修改的潜在可能 | • 要嵌入被监控程序中，修改被监控程序时容易引进错误<br>对策：探测器代码尽量短<br>• 不是分离进程，不易被禁止或修改 |
| 可实现性、可使用性和可维护性 | 好<br>• 探测器程序与被监控程序分离<br>• 从主机上进行修改、添加或删除等较容易<br>• 可以利用任何合适的编程语言 | 差<br>• 需要集成到被监控程序中，难度较大<br>• 需要使用与被监控程序相同的编程语言<br>• 设计要求高，修改和升级难度大 |
| 开销 | 大<br>数据生成和使用之间存在时延 | 小<br>• 数据的产生和使用之间的时延小<br>• 不是分离进程，避免了创建进程的主机开销 |
| 完备性 | 差<br>• 只能从“外面”监控程序<br>• 只能访问外部可以获得的数据，获取能力有限 | 好<br>• 可以放置在被监控程序的任何地方<br>• 可以访问被监控程序中的任何信息 |
| 正确性 | 只能根据可获得的数据作出基于经验的猜测 | 较好 |

入侵检测系统的数据分析技术依检测目标和数据属性，分为异常发现技术和模式发现技术两大类。最近几年还出现了一些通用的技术。下面分别进行介绍。

1）异常发现技术

异常发现技术用在基于异常检测的 IDS 中。如图 4.40 所示，在这类系统中，观测到的不是已知的入侵行为，而是所监视通信系统中的异常现象。如果建立了系统的正常行为轨迹，则在理论上就可以把所有与正常轨迹不同的系统状态视为可疑企图。由于正常情况具有一定的范围，因此正确地选择异常阈值和特征，决定何种程度才是异常，是异常发现技术的关键。

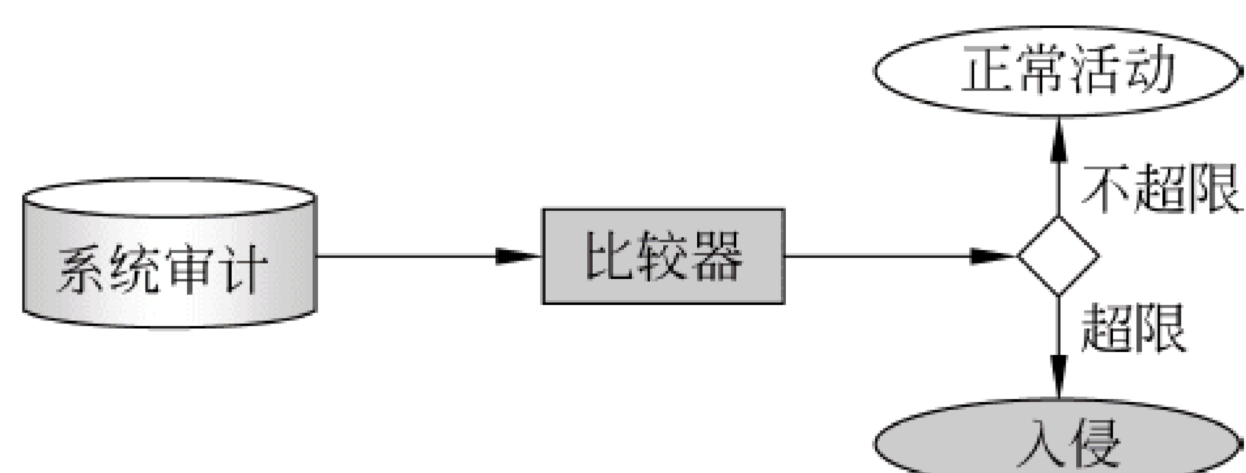


图 4.40　异常检测模型

异常检测只能检测出那些与正常过程具有较大偏差的行为。由于对各种网络环境的适应性较弱，且缺乏精确的判定准则，异常检测有可能出现虚报现象。

异常发现技术如表 4.11 所示。其中，自学习系统通过学习事例构建正常行为模型，又可分为时序和非时序两种；可编程系统需要通过程序测定异常事件，让用户知道哪些是足以破坏系统安全的异常行为，又可分为描述统计和缺省否定两类。

2）模式发现技术

模式发现又称为特征检测或滥用检测。如图 4.41 所示，它们是基于已知系统缺陷和入侵模式，即事先定义了一些非法行为，然后将观察现象与之比较做出判断。这种技术可以准确地检测具有某些特征的攻击，但是由于过度依赖实现定义好的安全策略，而无法检测

表 4.11　异常发现技术

| 类　型 | | 方　法 | 系 统 名 称 |
| --- | --- | --- | --- |
| 自学习型 | 非时序 | 规则建模 | Wisdom & Sense |
| | | 描述统计 | IDES、NIDES、EMERRALD、JiNao、Haystack |
| | 时序 | 人工神经网络 | Hyperview |
| 可编程型 | 描述统计 | 简单统计 | MIDAS、NADIR、Haystack |
| | | 基于简单规则 | NSM |
| | | 门限 | Computer-watch |
| | 缺省否认 | 状态序列建模 | DPEM、Janus、Bro |

系统未知的攻击行为，因而可能产生漏报。

模式发现技术对确知的决策规则通过编程实现，常用的技术有如下 4 种。

（1）状态建模：将入侵行为表示成许多个不同的状态。如果在观察某个可疑行为期间，所有状态都存在，则判定为恶意入侵。状态建模从本质上来讲是时间序列模型，可以再细分为状态转换和 Petri 网，前者将入侵行为的所有状态形成一个简单的遍历链，后者将所有的状态构成一个更广义的树形结构的 Petri 网。

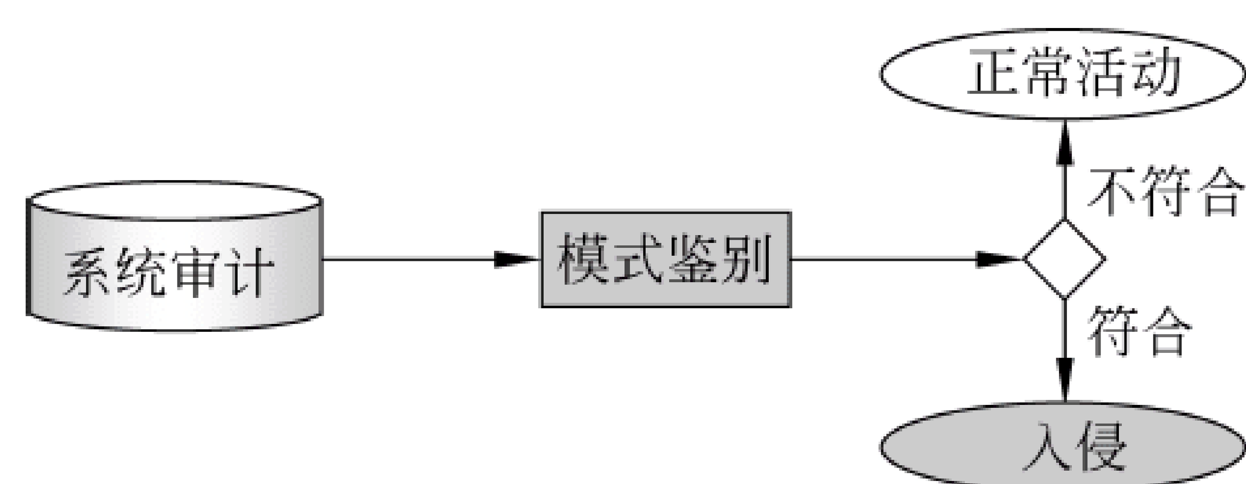



图 4.41　模式发现模型

（2）串匹配：通过对系统之间传输的或系统自身产生的文本进行子串匹配实现。该方法灵活性差，但易于理解，目前有很多高效的算法，其执行速度很快。

（3）专家系统：可以在给定入侵行为描述规则的情况下，对系统的安全状态进行推理。一般情况下，专家系统的检测能力强大，灵活性也很高，但计算成本较高，通常以降低执行速度为代价。

（4）基于简单规则：类似于专家系统，但相对简单一些，执行速度快。

3）混合检测

近几年来，混合检测日益受到人们的重视。这类检测在做出决策之前，既分析系统的正常行为，又观察可疑的入侵行为，所以判断更全面、准确、可靠。它通常根据系统的正常数据流背景来检测入侵行为，故也有人称其为“启发式特征检测”。属于这类检测的技术有以下一些。

（1）人工免疫方法。

（2）遗传算法。

（3）数据挖掘。

⋮

**4. 入侵检测系统的特征库**

IDS 要有效地捕捉入侵行为，必须拥有一个强大的入侵特征（signature）数据库，这就如同公安部门必须拥有健全的罪犯信息库一样。

IDS 中的特征是指用于判别通信信息种类的样板数据，通常分为多种，以下是一些典型情况及其识别方法。

（1）来自保留 IP 地址的连接企图：可通过检查 IP 报头（IP header）的来源地址识别。

（2）带有非法 TCP 标志联合物的数据包：可通过 TCP 报头中的标志集与已知正确和错误标记联合物的不同点来识别。

（3）含有特殊病毒信息的 E-mail：可通过对比每封 E-mail 的主题信息和病态 E-mail 的主题信息来识别，或者通过搜索特定名字的外延来识别。

（4）查询负载中的 DNS 缓冲区溢出企图：可通过解析 DNS 域及检查每个域的长度来识别。另外一个方法是在负载中搜索“壳代码利用”（exploit shellcode）的序列代码组合。

（5）对 POP3 服务器大量发出同一命令而导致 DoS 攻击：通过跟踪记录某个命令连续发出的次数，看看是否超过了预设上限，而发出报警信息。

（6）未登录情况下使用文件和目录命令对 FTP 服务器的文件访问攻击：通过创建具备状态跟踪的特征样板以监视成功登录的 FTP 对话，发现未经验证却发出命令的入侵企图。

显然，特征的涵盖范围很广，有简单的报头域数值，有高度复杂的连接状态跟踪，有扩展的协议分析。

此外，不同的 IDS 产品具有的特征功能也有所差异。例如，有些网络 IDS 系统只允许很少地定制存在的特征数据或者编写需要的特征数据，另外一些则允许在很宽的范围内定制或编写特征数据，甚至可以是任意一个特征；一些 IDS 系统，只能检查确定的报头或负载数值，另外一些则可以获取任何信息包的任何位置的数据。

### 4.4.3 响应与报警策略

**1. 响应**

早期的入侵检测系统的研究和设计把主要精力放在对系统的监控和分析上，而把响应的工作交给用户完成。现在的入侵检测系统都提供响应模块，并提供主动响应和被动响应两种响应方式。一个好的入侵检测系统应该让用户能够裁减定制其响应机制，以符合特定的需求环境。

1）主动响应

在主动响应系统中，系统将自动或以用户设置的方式阻断攻击过程或以其他方式影响攻击过程，通常可以选择的措施如下。

（1）针对入侵者采取的措施。

（2）修正系统。

（3）收集更详细的信息。

2）被动响应

在被动响应系统中，系统只报告和记录发生的事件。

**2．报警**

检测到入侵行为需要报警。具体报警的内容和方式需要根据整个网络的环境和安全需要确定。例如：

（1）对一般性服务企业，报警集中在已知的有威胁的攻击行为上。

（2）对关键性服务企业，需要将尽可能多的报警记录下来并对部分认定的报警进行实时反馈。

### 4.4.4　入侵检测器的部署与设置

入侵检测器部署的位置直接影响入侵检测系统的工作性能。在规划一个入侵检测系统时，首先要考虑入侵检测器的部署位置。显然，在基于网络的入侵检测系统中和在基于主机的入侵检测系统中，部署的策略不同。

**1．在基于网络的入侵检测系统中部署入侵检测器**

基于网络的入侵检测系统主要检测网络数据报文，因此一般将检测器部署在靠近防火墙的地方。具体做法有如图 4.42 所示的几个位置。

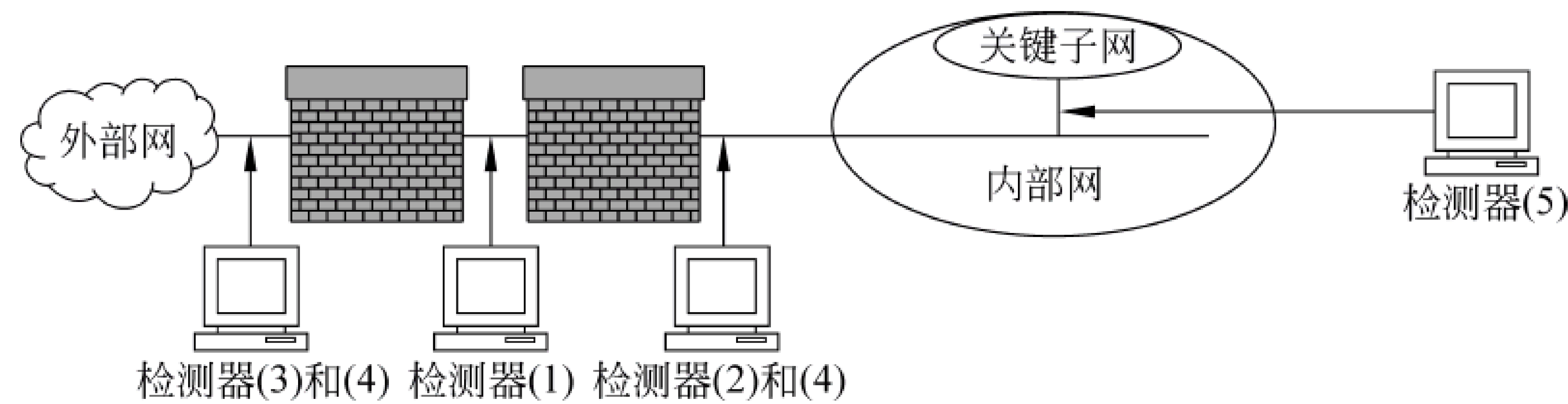



图 4.42　基于网络的入侵检测器的部署

1）DMZ 区

在这里，可以检测到的攻击行为是所有针对向外提供服务的服务器的攻击。由于 DMZ 中的服务器是外部可见的，因此在这里检测最为需要。同时，由于 DMZ 中的服务器有限，所以针对这些服务器的检测可以使入侵检测器发挥最大优势。但是，在 DMZ 中，检测器会暴露在外部而失去保护，容易遭受攻击，导致无法工作。

2）内网主干（防火墙内侧）

将检测器放到防火墙的内侧，有如下几点好处。

（1）检测器比放在 DMZ 中安全。

（2）所检测到的都是已经渗透过防火墙的攻击行为。从中可以有效地发现防火墙配置的失误。

（3）可以检测到内部可信用户的越权行为。

（4）由于受干扰的机会少，报警概率也小。

3）外网入口（防火墙外侧）

这种部署的优势如下。

（1）可以对针对目标网络的攻击进行计数，并记录最为原始的数据包。

（2）可以记录针对目标网络的攻击类型。

但是，检测器部署在外网入口不能定位攻击的源地址和目的地址，系统管理员在处理攻击行为上也有难度。

4）在防火墙的内外都放置

这种位置既可以检测到内部攻击，又可以检测到外部攻击，并且无须猜测攻击是否穿越防火墙。但是，这种部署的开销较大。在经费充足的情况下是最理想的选择。

5）关键子网

这个位置可以检测到对系统关键部位的攻击，将有限的资源用在最值得保护的地方，获得最大效益/投资比。

**2. 在基于主机的入侵检测系统中部署入侵检测器**

基于主机的入侵检测系统通常是一个程序。在基于网络的入侵检测器的部署和配置完成后，基于主机的入侵检测将部署在最重要、最需要保护的主机上。

**3. 入侵检测系统的设置**

网络安全需要各个安全设备的协同工作和正确设置。由于入侵检测系统位于网络体系中的高层，高层应用的多样性导致了入侵检测系统分析的复杂性和对计算资源的高需求。在这种情形下，对入侵检测设备进行合理的优化设置，可以使入侵检测系统更有效地运行。

图 4.43 是入侵检测系统设置的基本过程。可以看出，入侵检测系统的设置需要经过多次回溯，反复调整。

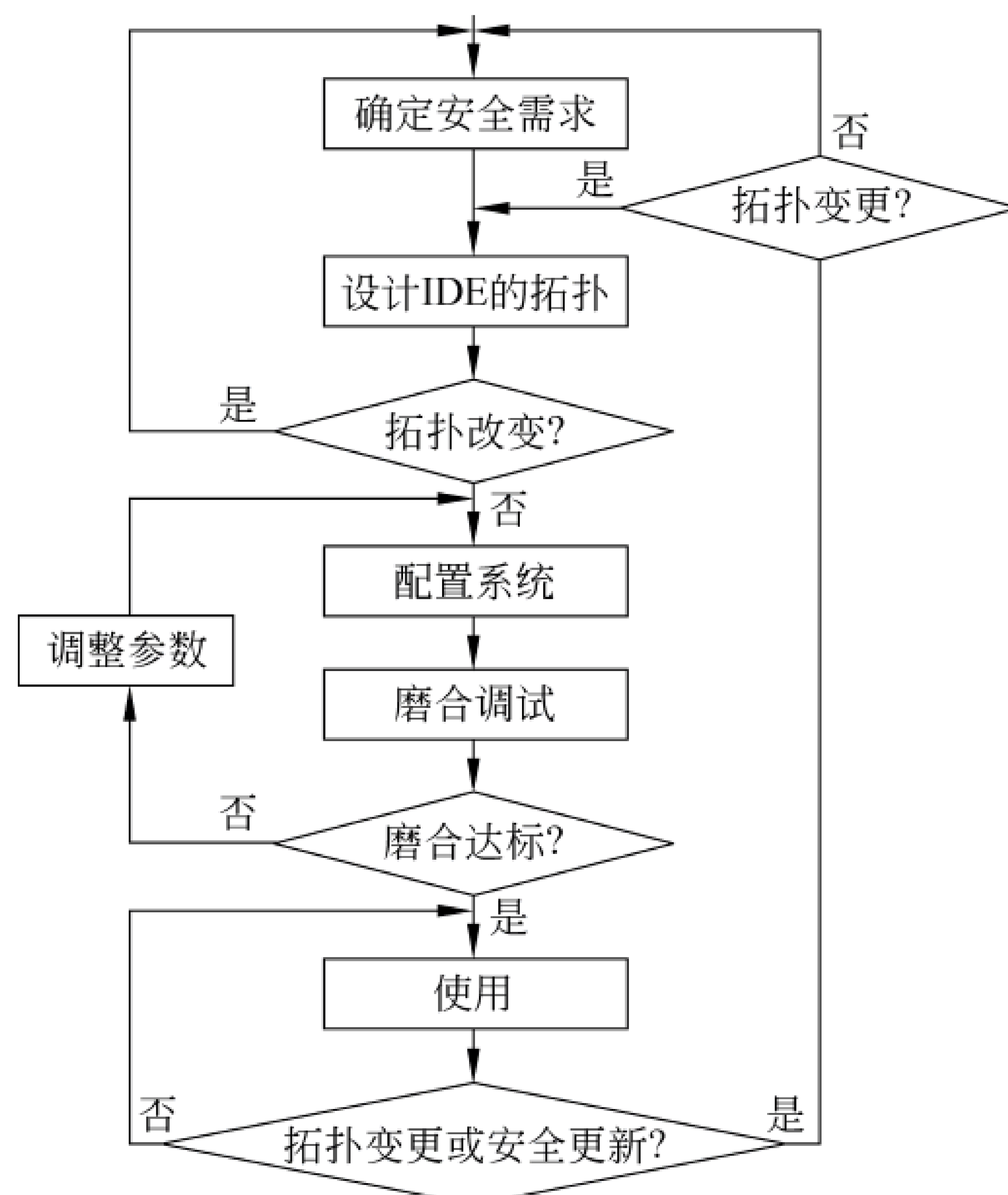



图 4.43　入侵检测系统设置的基本过程

### 4. 报警策略

检测到入侵行为时需要报警。具体报警的内容和方式需要根据整个网络的环境和安全需要确定。例如：

（1）对一般性服务企业，报警集中在已知的有威胁的攻击行为上。

（2）对关键性服务企业，需要将尽可能多的报警记录下来并对部分认定的报警进行实时反馈。

## 4.5 网 络 诱 骗

防火墙和入侵检测都是被动防御技术，而网络诱骗是一种主动防御技术。表 4.12 为两种防御系统之间的工作特点比较。

**表 4.12 主动防御与被动防御的比较**

| 项　目 | 被动防御系统 | 主动防御系统 |
| --- | --- | --- |
| 主动性 | 被动地“守株待兔”式防御 | 主动跟踪攻击者 |
| 攻防方式 | 攻击者主动选择攻击目标，可随意攻击 | 事先掌握攻击者的行为，进行跟踪，有效制止攻击者的破坏行为 |
| 对攻击者的威慑 | 对攻击方不构成威胁 | 能对攻击者造成威胁和损害：<br>• 诱惑黑客攻击虚假网络而忽视真正的网络<br>• 加重黑客的工作量，消耗其资源，让系统管理员有足够时间响应<br>• 收集黑客信息和企图，以便系统进行安全防护和检测<br>• 为起诉留下证据 |

### 4.5.1 蜜罐主机技术

网络诱骗技术的核心是蜜罐（honey pot）。它是运行在 Internet 上的充满诱惑力的计算机系统。这种计算机系统有如下一些特点。

（1）蜜罐是一个包含漏洞的诱骗系统，它通过模拟一个或多个易受攻击的主机，给攻击者提供一个容易攻击的目标。

（2）蜜罐不向外界提供真正有价值的服务。

（3）所有与蜜罐的连接尝试都被视为可疑的连接。

这样，蜜罐就可以实现如下目的。

（1）引诱攻击，拖延对真正有价值目标的攻击。

（2）消耗攻击者的时间，以便收集信息，获取证据。

下面介绍蜜罐的 3 种主要形式。

#### 1. 空系统

空系统是一种没有任何虚假和模拟的环境的完全真实的计算机系统，有真实的操作系统和应用程序，也有真实的漏洞。这是一种简单的蜜罐主机。

但是，空系统（以及模拟系统）会很快被攻击者发现，因为他们会发现这不是期待的

目标。

**2．镜像系统**

建立一些提供 Internet 服务的服务器镜像系统，会使攻击者感到真实，也就更具有欺骗性。另一方面，由于是镜像系统，所以比较安全。

**3．虚拟系统**

虚拟系统是在一台真实的物理机器上运行一些仿真软件，模拟出多台虚拟机，构建多个蜜罐主机。这种虚拟系统不但逼真，而且成本较低，资源利用率较高。此外，即使攻击成功，也不会威胁宿主操作系统的安全。

### 4.5.2 蜜网技术

蜜网（honey net）技术也称为陷阱网络技术。它由多个蜜罐主机、路由器、防火墙、IDS、审计系统等组成，为攻击者制造一个攻击环境，供防御者研究攻击者的攻击行为。

**1．第一代蜜网**

图 4.44 为第一代蜜网的结构。

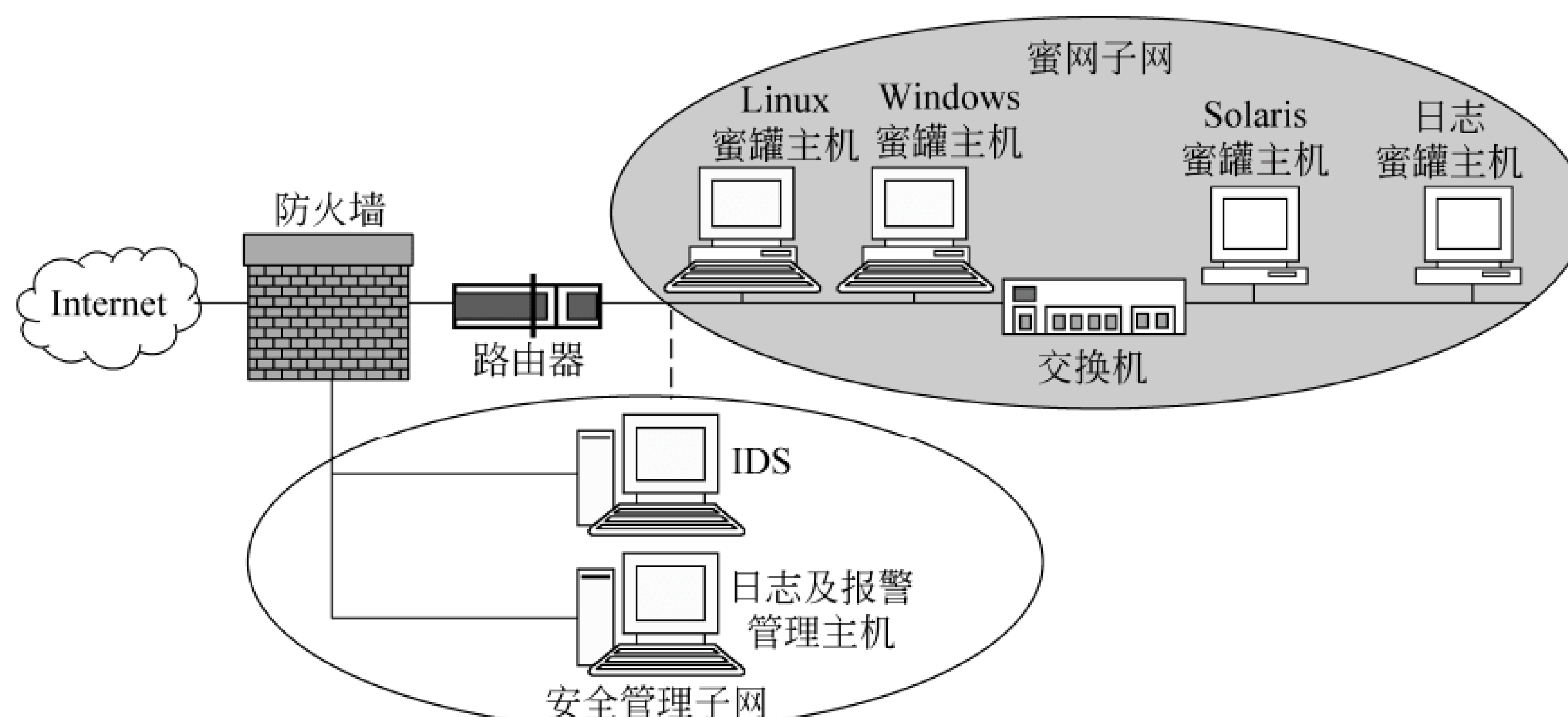



图 4.44 第一代蜜网的结构

下面对其中各部件的作用加以介绍。

（1）防火墙：用于隔离内网和外网，防止入侵者以蜜网作为跳板攻击其他系统。其配置规则为：不限制外网对蜜网的访问，但需要对蜜罐主机对外的连接予以控制，包括：

① 限制对外连接的目的地。

② 限制蜜罐主机主动对外连接。

③ 限制对外连接的协议。

⋮

（2）路由器：放在防火墙与蜜网之间，利用路由器具有的控制功能来弥补防火墙的不足，例如，防止地址欺骗攻击和 DoS 攻击等。

（3）IDS：是蜜网中的数据捕获设备，用于检测和记录网络中可疑的通信连接，在发现

可疑的网络活动时报警。

### 2. 第二代蜜网

图 4.45 为第二代蜜网的结构图。

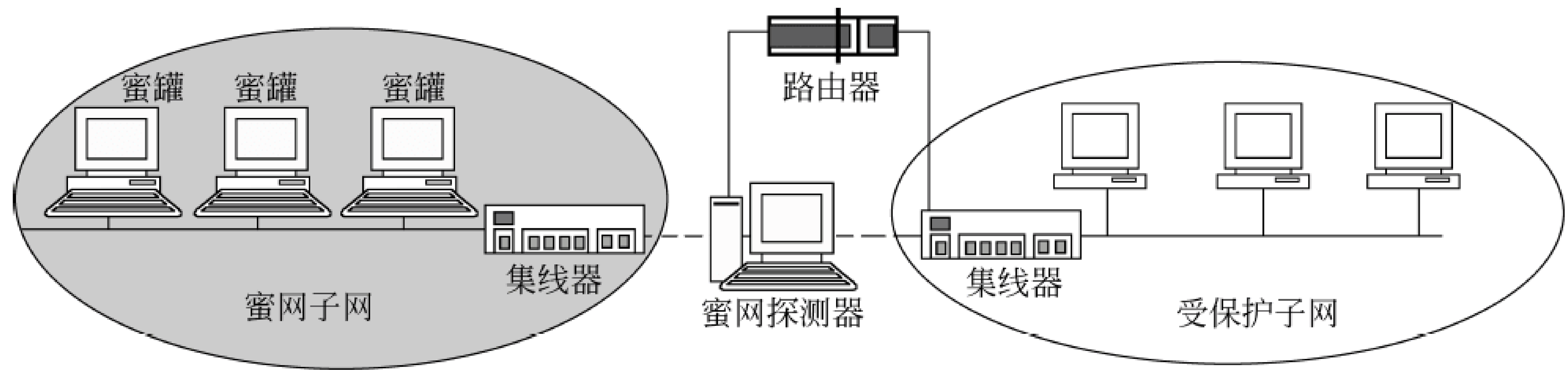



图 4.45　第二代蜜网的结构

第二代蜜网技术将数据控制和数据捕获集中到蜜网探测器中进行。这样带来的好处如下。

（1）便于安装和管理。

（2）隐蔽性更强。

（3）可以监控非授权活动。

（4）可以采取积极的响应方法限制非法活动的效果，如修改攻击代码字节，使攻击失效等。

### 3. 第三代蜜网

第三代蜜网是目前正在开发的蜜网技术。它是建立在物理设备上的分布式虚拟系统，如图 4.46 所示，这样就把蜜罐、数据控制、数据捕获和数据记录等都集中到一台物理设备上。




图 4.46　第三代蜜网结构

## 4.5.3　常见网络诱骗工具及产品

### 1. 蜜罐实现工具

1）winetd

winetd 是一个在 Windows 上实现蜜罐的简单工具。它安装简单，界面友好，适合初学者使用；缺点是过于简单，并不能真正诱骗攻击者进入。

2）DTK

DTK（Deception Tool Kit，可以从 http://all.net/dtk/网站下载）是用 C 语言和 Perl 脚本语言写成的一种蜜罐工具软件，能在支持 C 语言和 Perl 的系统（UNIX）上运行。它能够监听 HTTP、FTP 和 Telnet 等常用服务器所使用的端口，模拟标准服务器对接收到的请求所做

出的响应，还可以模拟多种常见的系统漏洞。其不足之处是：模拟不太逼真，构建过程麻烦。

3）Honeyd

Honeyd（可以从 http://www.citi.umich.edu/u/provos/honeyd 网站下载）是一个专用的蜜罐构建软件，可以虚拟多种主机，配置运行不同的服务和操作系统。

**2. 蜜网实现工具**

（1）数据控制工具：Jptable、snort_inline。
（2）数据捕获工具：Termlog、Sebek2、snort、Comlog。
（3）数据收集工具：Obfugator。
（4）数据分析工具：Privmsg、TASK、WinInterrogate。
以上工具可以从下面的网址下载：http://project.honeynet.org/。

# 习　题

## 一、选择题

1. 防火墙用于将Internet和内部网络隔离，是________。
   A. 防止Internet火灾的硬件设施　B. 网络安全和信息安全的软件和硬件设施
   C. 保护线路不受破坏的软件和硬件设施　D. 起抗电磁干扰作用的硬件设施
2. 防火墙最主要被部署在________位置。
   A. 网络边界　B. 骨干线路　C. 重要服务器旁　D. 桌面终端
3. 下列关于防火墙的说法中错误的是________。
   A. 防火墙工作在网络层　B. 防火墙对IP数据包进行分析和过滤
   C. 防火墙是重要的边界保护机制　D. 部署防火墙，就解决了网络安全问题
4. 在一个企业网中，防火墙应该是________的一部分，构建防火墙时首先要考虑其保护的范围。
   A. 安全技术　B. 安全设置　C. 局部安全策略　D. 全局安全策略
5. 一般而言，Internet防火墙建立在一个网络的________。
   A. 内部子网之间传送信息的中枢　B. 每个子网的内部
   C. 内部网络与外部网络的交叉点　D. 部分内部网络与外部网络的结合处
6. 包过滤型防火墙从原理上看是基于________进行数据包分析的技术。
   A. 物理层　B. 数据链路层　C. 网络层　D. 应用层
7. 对非军事DMZ而言，正确的解释是________。
   A. DMZ是一个真正可信的网络部分
   B. DMZ网络访问控制策略决定允许或禁止进入DMZ通信
   C. 允许外部用户访问DMZ系统上合适的服务
   D. 以上3项都对
8. 对动态网络地址交换（NAT），不正确的说法是________。

A. 将很多内部地址映射到单个真实地址　　B. 外部网络地址和内部地址一对一地映射

C. 每个连接使用一个端口　　D. 最多可有64 000个同时的动态NAT连接

9. 以下________不是包过滤防火墙主要过滤的内容。

A. 源IP地址　　B. 目的IP地址

C. TCP源端口和目的端口　　D. 时间

10. 在被屏蔽的主机体系中，堡垒主机位于________中，所有的外部连接都经过滤路由器到它上面去。

A. 内部网络　　B. 周边网络　　C. 外部网络　　D. 自由连接

11. 外部数据包经过过滤路由只能阻止________的唯一IP欺骗。

A. 内部主机伪装成外部主机IP　　B. 内部主机伪装成内部主机IP

C. 外部主机伪装成外部主机IP　　D. 外部主机伪装成内部主机IP

12. IPSec协议工作在网络的________。

A. 数据链路层　　B. 网络层　　C. 应用层　　D. 传输层

13. IPSec协议中涉及密钥管理的重要协议是________。

A. IKE　　B. AH　　C. ESP　　D. SSL

14. SSL产生会话密钥的方式是________。

A. 从密钥管理数据库中请求获得

B. 每一台客户机分配一个密钥的方式

C. 随机由客户机产生并加密后通知服务器

D. 由服务器产生并分配给客户机

15. 传输层保护的网络采用的主要技术是建立在________基础上的________。

A. 可靠的传输服务　安全套接字层（SSL）协议

B. 不可靠的传输服务　S-HTTP协议

C. 可靠的传输服务　S-HTTP协议

D. 不可靠的传输服务　安全套接字层（SSL）协议

16. 主要用于加密机制的协议是________。

A. HTTP　　B. FTP　　C. Telnet　　D. SSL

17. 通常所说的移动VPN是指________。

A. Access VPN　　B. Intranet VPN　　C. Extranet VPN　　D. 以上均不是

18. 以下属于第二层的VPN隧道协议有________。

A. IPSec　　B. PPTP　　C. GRE　　D. 以上均不是

19. 将公司与外部供应商、客户及其他利益相关群体相连接的是________。

A. 内联网VPN　　B. 外联网VPN　　C. 远程接入VPN　　D. 无线VPN

20. 以下不属于隧道协议的是________。

A. PPTP　　B. L2TP　　C. TCP/IP　　D. IPSec

21. 以下不属于VPN核心技术的是________。

A. 隧道技术　　B. 身份认证　　C. 日志记录　　D. 访问控制

22. ________通过一个拥有与专用网络相同策略的共享基础设施，提供对企业内部网或外部网的远程访问。

A. Access VPN　　B. Intranet VPN　　C. Extranet VPN　　D. Internet VPN

23. L2TP隧道在两端的VPN服务器之间采用________来验证对方的身份。

A. SSL　　B. 数字证书　　C. Kerberos　　D. 口令握手协议CHAP

24. 入侵检测的基本方法是________。

A. 基于用户行为概率统计模型的方法　　B. 基于神经网络的方法

C. 基于专家系统的方法　　D. 以上都正确

25. 关于入侵检测技术，下列描述中错误的是________。

A. 入侵检测系统不对系统或网络造成任何影响

B. 审计数据或系统日志信息是入侵检测系统的一项主要信息来源

C. 入侵检测信息的统计分析有利于检测到未知的入侵和更为复杂的入侵

D. 基于网络的入侵检测系统无法检查加密的数据流

## 二、填空题

1. 防火墙位于________之间，一端是________，另一端是________。

2. 防火墙系统的主要体系结构有________体系结构、________体系结构和________体系结构。

3. 按检测的监控位置划分，入侵检测系统可分为基于________的入侵检测系统、基于________的入侵检测系统和________入侵检测系统。

4. ________被定义为通过公用网络建立的一个临时的、安全的连接，是一条穿过公用网络的安全、稳定的通道。

## 三、问答题

1. 在组建Intranet时，防火墙是必需的吗？为什么？

2. 试述一个防火墙产品应具备哪些基本功能。

3. 下面是选择防火墙时应考虑的一些因素，请按你的理解，将它们按重要性排序。

（1）被保护网络受威胁的程度。

（2）受到入侵，网络的损失程度。

（3）网络管理员的经验。

（4）被保护网络的已有安全措施。

（5）网络需求的发展。

（6）防火墙自身管理的难易度。

（7）防火墙自身的安全性。

4. 列举更多的防火墙系统结构，最好有自己的创意。

5. 查找资料，叙述防火墙测试的内容和方法。

6. 查找资料，叙述防火墙选型的基本原则和具体标准。

7. 简述攻击防火墙的主要手段。

8. 查找资料，简述目前国内外防火墙技术发展的现状和自己对防火墙的未来的设想。

9. 收集资料，对当前常用的防火墙产品进行分析比较，详细描述其中的3种防火墙产品的用法以及升级方法。

10. 浏览最热门的3个防火墙技术网站，综述目前关于防火墙讨论的热点问题。

11. 有一个内部网（192.168.20.0）只与某一台外部主机（172.1652.55）交换数据。写出位于它们之间的数据包过滤规则。

12. 比较包过滤、网络地址转换和代理技术的特点以及适用的环境。

13. 简述国内外物理隔离技术的现状和发展趋势。

14. 浏览网站，列举国内有关物理隔离设备的厂家及其产品的特点。

15. 收集国内外有关网络隔离技术的网站信息，简要说明各网站的特点。

16. 收集国内外有关网络隔离技术的最新动态。

17. 简述VPN使用了哪些主要技术。

18. 综述有关入侵检测技术的各种定义。

19. 入侵检测系统有哪些可以利用的数据源？

20. 试构造一个网络数据包的截获程序。

21. 试述入侵检测系统的工作原理。

22. 收集资料，对国内外主要基于网络的入侵检测产品进行比较。

23. 收集资料，对国内外主要基于主机的入侵检测产品进行比较。

24. 分析入侵检测系统的不足和发展趋势。

25. 入侵检测技术与法律有什么关系？

26. 简述蜜罐技术的特殊用途。

27. 用下载的蜜罐工具构造一个简单的蜜罐系统。

28. 收集国内外有关入侵检测、网络诱骗的最新动态。

# 第 5 章　信息系统安全管理

信息系统是复杂的系统，其安全的运行不仅与技术有关，还涉及人和制度。因此，从用户的角度看，更需要强调三分技术、七分管理。

经过多年的实践和研究，各个国家以及组织，都已形成完善的信息安全管理体系和方法。本章介绍信息系统管理中的一些主要环节。

## 5.1　信息系统应急响应

“智者千虑，必有一失。”尽管人们已经为信息系统的防护开发了许多技术，但是很难没有一点疏漏，何况入侵者也是一些技术高手。

系统遭受到一次入侵，就面临一次灾难。这些影响信息系统安全的不正当行为就称为事件。事件响应就是事件发生后所采取的措施和行动。信息系统的脆弱，加上入侵技术的不断进化，使得入侵不可避免。因此，安全事件响应就成为一个与防火墙技术、入侵检测技术等同样重要的安全保障策略和手段。

1988 年，莫里斯蠕虫以迅雷不及掩耳之势肆虐互联网，招致上千台计算机系统的崩溃，造成了以千万美元计的损失。这突如其来的灾难给人们敲响了警钟：面对人类对信息系统依赖程度的不断增强，对付入侵不仅需要防御，还要能够在事件发生后进行紧急处理和援助。1989 年，在美国国防部的资助下，计算机紧急响应组/呼叫中心（Computer Emergency Team/Call Center，CERT/CC）成立。从此紧急响应被摆到了人们的议事桌上。

一般说来，每个使用信息系统的组织都应当有一套应急响应机制。该机制应包括 4 个基本环节。

（1）信息系统应急响应组织。

（2）信息系统安全保护制度。

（3）信息系统应急预案。

（4）信息系统应急演练。

### 5.1.1　应急响应组织

应急响应组织的主要工作如下。

（1）安全事件与软件安全缺陷分析研究。

（2）安全知识库（包括漏洞知识、入侵检测等）开发与管理。

（3）安全管理和应急知识的教育与培训。

（4）发布安全信息（如系统漏洞与补丁、病毒警告等）。

（5）安全事件紧急处理。

应急响应组织包括应急保障领导小组和应急技术保障小组。应急保障领导小组的主要

职责是领导与协调突发事件与自然灾害的应急处理。应急技术保障小组主要解决安全事件的技术问题，如物理实体和环境安全技术、网络通信技术、系统平台技术、应用系统技术等。当然其中有些工作也可以进行服务外包。

### 5.1.2 信息系统安全保护制度

信息系统安全保护制度主要包括如下 4 个方面的内容。

（1）信息系统安全责任制度。

（2）信息系统安全检测制度。

（3）信息系统安全报告制度。

（4）信息安全行为规范。

#### 1. 信息系统安全责任制度

信息系统安全责任制度包括如下几个方面。

1）责任到人

例如，《广东省计算机信息系统安全保护管理规定》（经 2003 年 3 月 31 日广东省人民政府第十届 4 次常务会议通过，2003 年 4 月 8 日广东省人民政府令第 81 号发布，自 2003 年 6 月 1 日起施行）要求："计算机信息系统使用单位应当确定计算机安全管理责任人，建立健全安全保护制度，落实安全保护技术措施，保障本单位计算机信息系统安全，并协助公安机关做好安全保护管理工作。"

2）确定信息系统安全保护的重点部门

例如，对于一个地区，应当把下列计算机信息系统使用单位为重点安全保护单位。

（1）县级以上国家机关、国防单位。

（2）银行、证券、能源、交通、邮电通信单位。

（3）国家及省重点科研、教育单位。

（4）国有大中型企业。

（5）互联单位、接入单位及重点网站。

（6）向公众提供上网服务的场所。

3）建立信息系统安全从业资质、证书制度

例如，《广东省计算机信息系统安全保护管理规定》要求：

（1）重点安全保护单位计算机安全管理责任人和信息审查员应当参加县级以上人民政府公安机关认可的安全技术培训，并取得安全技术培训合格证书。

（2）申请安全服务资质，应当具备以下条件。

① 取得相应经营范围的营业执照。

② 取得安全技术培训合格证书的专业技术人员不少于 10 人，其中大学本科以上学历的人员所占比例不少于 70%。

③ 负责安全服务工作的管理人员应当具有两年以上从事计算机信息系统安全技术领

域企业管理工作经历，并取得安全技术培训合格证书。

④ 有与其从事的安全服务业务相适应的技术装备。

⑤ 有与其从事的安全服务业务相适应的组织管理制度。

⑥ 法律法规规定的其他条件。

4）信息安全奖惩制度

例如，《广东省计算机信息系统安全保护管理规定》（2003）要求："计算机信息系统使用单位应当将计算机信息系统安全保护工作纳入内部检查、考核、评比内容。对在工作中成绩突出的部门和个人，应当给予表彰奖励。对未依法履行安全保护职责或违反本单位安全保护制度的，应当依照有关规定对责任人员给予行政处分。"

### 2. 信息系统安全检测制度

信息系统维护与管理人员要定期对信息系统进行安全检测，防患于未然，包括定期检查下列内容。

（1）系统重要部分的冗余或备份措施。

（2）计算机病毒防治措施。

（3）网络攻击防范、追踪措施。

（4）安全审计和预警措施。

（5）系统运行和用户使用日志记录保存 60 日以上措施。

（6）记录用户主叫电话号码和网络地址的措施。

（7）身份登记和识别确认措施。

（8）信息群发限制和有害数据防治措施。

### 3. 信息系统安全报告制度

要建立信息系统安全报告制度，按照有关规定，向有关部门报告信息系统安全运行情况以及有关事件。例如，《广东省计算机信息系统安全保护管理规定》（2003）要求"对计算机信息系统中发生的重大安全事故，使用单位应当采取应急措施，保留有关原始记录，在 24 小时内向当地县级以上人民政府公安机关报告"。

### 4. 信息安全行为规范

通常要求任何单位和个人不得利用计算机信息系统从事下列行为。

（1）制作、复制、查阅、传播有害信息。

（2）侵犯他人隐私，窃取他人账号，假冒他人名义发送信息，或者向他人发送垃圾信息。

（3）以营利或者非正常使用为目的，未经允许向第三方公开他人电子邮箱地址。

（4）未经允许修改、删除、增加、破坏计算机信息系统的功能、程序及数据。

（5）危害计算机信息系统安全的其他行为。

### 5.1.3　信息系统应急预案

#### 1. 应急预案及基本内容

应急预案是指根据不同的突发紧急事件类型和意外情形预先制定的处理方案。应急预案一般要包括如下内容。

（1）执行应急预案的人员（姓名、住址、电话号码以及有关职能部门的联系方法）。

（2）系统紧急事件类型及处理措施的详细说明。

（3）应急处理的具体步骤和操作顺序。

#### 2. 常见安全事件举例

应急预案要根据安全事件的类型进行对应的处理。不同的组织中，信息系统安全事件的常见类型有所不同，工作人员应当根据行业性质和地域状况进行具体分析。下面提供一些常见的安全事件类型供参考。

（1）物理实体及环境类安全事件，如意外停电、物理设备丢失、火灾、水灾等。

（2）网络通信类安全事件，如网络蠕虫侵害等。

（3）主机系统类安全事件，如计算机病毒、口令丢失等。

（4）应用系统类安全事件，如客户信息丢失等。

#### 3. 应急事件处理的基本流程

1）安全事件报警

值班人员发现紧急情况，要及时报告。报告要对安全事件进行准确描述并作书面记录。按照安全事件的类型，安全事件呈报条例应依次报告：值班人员、应急工作组长、应急领导小组。如果想进行任何类型的跟踪调查或者起诉入侵者，应先跟管理人员和法律顾问商量，然后通知有关执法机构。一定要记住，除非执法部门的参与，否则对入侵者进行的一切跟踪都可能是非法的。

同时，还应通知有关人员，交换相关信息，必要时可以获得援助。

2）信息安全紧急事件认定

信息系统发生下列事件之一，应视为紧急事件。

（1）信息系统硬件受到破坏性攻击，不能正常发挥其部分或全部功能。

（2）信息系统软件受到破坏性攻击，不能正常发挥其部分或全部功能。

（3）信息系统受到恶意程序攻击，局部或全部数据或功能受到损坏，或工作效率急剧降低。

（4）相关物理设备受到人为和自然灾害破坏，无法正常工作。

（5）出现意外停电而又无后备电源。

（6）关键岗位人员不能到位。

紧急事件发生后，应尽快确定安全事件的类型，以便启动相应的预案。

3）启动应急预案

（1）首先要能够找到应急预案。
（2）保护现场证据（如系统事件、处理者采取的行动、与外界的沟通等），避免灾害扩大。
（3）控制事态发展。

4）恢复系统

这部分内容将在 5.1.4 节中详细介绍。

5）应急工作总结

召开会议，分析问题和解决方法，具体可参考 ftp://ftp.isi.edu/in-notes/rfc2196.txt。
（1）总结教训。从记录中总结关于这起事故的教训，这有助于反思安全策略。
（2）计算事件的代价。计算事件代价有助于让组织认识到安全的重要性。
（3）改进安全策略。

6）撰写安全事件报告

安全事件报告的内容包括以下部分。
（1）安全事件发生的日期、时间。
（2）安全事件处理参加的人员。
（3）事件发现的途径。
（4）事件类型。
（5）事件涉及范围。
（6）现场记录。
（7）事件导致的损失和影响。
（8）事件处理过程。
（9）使用的技术和工具。
（10）经验和教训。

### 5.1.4 灾难恢复

灾难恢复是安全事件应急预案中特别重要的部分，从发现入侵的那一刻起，所有工作就都围绕它进行。

#### 1. 灾难恢复的内容

灾难恢复中应当包括如下几项内容。

1）与高层管理人员协商

系统恢复的步骤应当符合组织的安全预案。如果安全预案中没有描述，应当与管理人员协商，以便能从更高角度进行判断，并得到更多部门的支持和配合。

2）夺回系统控制权

为了夺回对被入侵系统的控制权，先需要将被入侵系统从网络上断开，包括拨号连接。

如果在恢复过程中，没有断开被侵入系统和网络的连接，入侵者就可能破坏所进行的恢复工作。

进行系统恢复也会丢失一些有用信息，如入侵者正在使用的扫描程序或监听进程。因此想要继续追踪入侵者时，可以先不夺回系统控制权，以免被入侵者发现。但是，也要采取其他一些措施，避免入侵蔓延。

3）复制一份被入侵系统的映像

在进行入侵分析之前，最好对被入侵系统进行备份（如使用 UNIX 命令 dd）。这个备份在恢复失败时非常有用。

4）入侵评估

入侵评估包括入侵风险评估、入侵路径分析、入侵类型确定和入侵涉及范围调查。下面介绍围绕这些工作进行的调查工作。

（1）详细审查系统日志文件和显示器输出，检查异常现象。

（2）入侵者遗留物分析，包括以下几方面。

① 检查入侵者对系统文件和配置文件的修改。

② 检查被修改的数据。

③ 检查入侵者留下的工具和数据。

④ 检查网络监听工具。

（3）其他，如网络的周边环境和涉及的远程站点。

5）清除后门

后门是入侵者为下次攻击设下的埋伏，包括修改了的配置文件、系统木马程序、修改了的系统内核等。

6）查阅 CERT 的安全建议、安全总结和供应商的安全提示

查阅 CERT 以往的安全建议和总结以及供应商的安全提示，一定要安装所有的安全补丁。

（1）CERT 安全建议：见 http://www.cert.org/advisories/。

（2）CERT 安全总结：见 http://www.cert.org/advisories/。

（3）供应商安全提示：见 ftp：//ftp.cert.org/pub/cert_bulletins/。

7）记录恢复过程中所有的步骤

毫不夸张地讲，记录恢复过程中采取的每一步措施都是非常重要的。恢复一个被入侵的系统是一件很麻烦的事，要耗费大量的时间，因此经常会使人做出一些草率的决定。记录恢复过程的每一步可以帮助自己避免做出草率的决定，还可以留作以后参考，也可能给法律调查提供帮助。

8）系统恢复

各种安全事件预案的执行都是为了使系统在事故后得以迅速恢复。对于服务器和数据库等特别重要的设备，则需要单独订立紧急恢复预案。

（1）操作系统恢复。

① 安装干净的操作系统版本。如果主机被入侵，就应当考虑系统中的任何东西都可能被攻击者修改过了，包括内核、二进制可执行文件、数据文件、正在运行的进程以及内存。通常，需要从发布介质上重装操作系统，然后在重新连接到网络上之前，安装所有的安全补丁，只有这样才会使系统不受后门和攻击者的影响。只找出并修补被攻击者利用的安全缺陷是不够的。

建议使用干净的备份程序备份整个系统。然后重装系统。

② 取消不必要的服务。只配置系统要提供的服务，取消那些没有必要的服务。检查并确信其配置文件没有脆弱性以及该服务是否可靠。通常，最保守的策略是取消所有的服务，只启动所需要的服务。

③ 安装供应商提供的所有补丁。建议安装所有的安全补丁，使系统能够抵御外来攻击，不被再次入侵，这是最重要的一步。

④ 安装所有需要的驱动程序。

⑤ 安装所有需要的服务软件包。

⑥ 安装备份软件及其修补程序。

显然，用手工进行服务器的恢复是非常麻烦的。如果能设计一种专门的软件包，可以生成存有服务器镜像文件的启动盘来恢复服务器，就便利多了。

（2）数据库系统的恢复。

数据库恢复是指通过技术手段，将保存在数据库中丢失的电子数据进行抢救和恢复的技术，其中包括：

① 数据文件恢复：把备份文件恢复到原来位置。

② 控制文件恢复：控制文件受损时，要将其恢复到原位重新启动。

③ 文件系统恢复：在大型操作系统中，可能会因介质受损导致文件系统被破坏。

数据库恢复的一般步骤如下。

① 将介质重新初始化。

② 重新创建文件系统。

③ 利用备份完整地恢复数据库中的数据。

④ 启动数据库系统。

（3）数据恢复。

谨慎使用备份数据。在从备份中恢复数据时，要确信备份主机没有被入侵。一定要记住，恢复过程可能会重新带来安全缺陷，被入侵者利用。例如，备份中的用户的 home 目录（UNIX/Linux 系统中有一个/home 目录，通常用来保存用户的文件，并且一个用户登录系统并进入后所处的位置就是/home 目录）、数据文件、hosts 文件（hosts 是一个没有扩展名的系统文件，可以用记事本等工具打开，其作用就是将一些常用的网址域名对应的 IP 地址建立一个关联“数据库”）中也许藏有特洛伊木马程序。

9）改变密码

在弥补了安全漏洞或者解决了配置问题以后，建议改变系统中所有账户的密码。

10）加固系统和网络的安全

（1）根据 CERT 的配置指南检查系统的安全性。

CERT 的 UNIX/NT 配置指南可以帮助用户检查系统中容易被入侵者利用的配置问题。具体可查阅以下两个网络资源：

http://www.cert.org/tech_tips/unix_configuration_guidelines.html

http://www.cert.org/tech_tips/win_configuration_guidelines.html

查阅安全工具文档可以参考 http://www.cert.org/tech_tips/security_tools.html。

（2）安装安全工具。在将系统连接到网络上之前，一定要安装所有选用的安全工具。同时，最好使用 Tripwire、aide 等工具对系统文件进行 MD5 校验，把校验码放到安全的地方，以便以后对系统进行检查。

（3）打开日志。启动日志（logging）/检查（auditing）/记账（accounting）程序，将它们设置到准确的级别，例如 sendmail 日志应该是 9 级或者更高。

要经常备份日志文件，或者将日志写到另外的计算机、一个只能增加的文件系统或者一个安全的日志主机。

（4）配置防火墙对网络进行防御。可以参考 http://www.cert.org/tech_tips/packet_filtering.html。

（5）重新连接到 Internet。完成以上步骤以后，就可以把系统重新连入 Internet 了。

应当注意，安全事件处理工作复杂，责任重大，至少应有两人参加。

**2. 灾难恢复级别**

2007 年 7 月，国务院信息化工作办公室下发了《信息系统灾难恢复规范》，并于 2007 年 11 月 1 日开始正式实施。这是中国灾难备份与恢复行业的第一个国家标准，也是各行各业进行灾备建设的重要参考性文件。GB/T 20988—2007《信息安全技术 信息系统灾难恢复规范》将灾难恢复能力划分为如图 5.1 所示的 6 级。

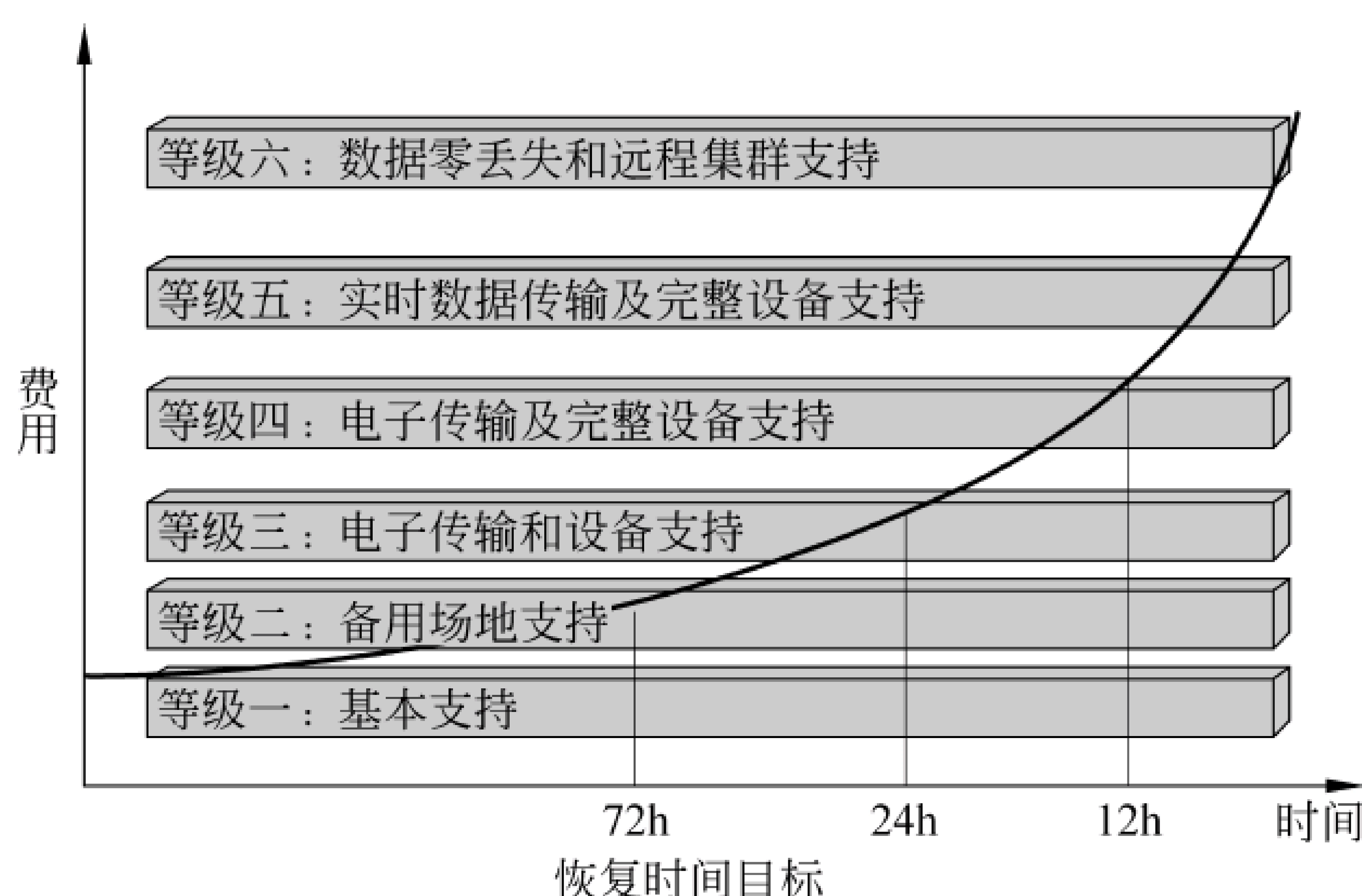



图 5.1 信息系统灾难恢复级别

等级一：基本支持。要求数据备份系统能够保证每周至少进行一次数据备份，备份介质能够提供场外存放。对于备用数据处理系统和备用网络系统，没有具体要求。

等级二：备用场地支持。在满足等级一的条件基础上，要求配备灾难恢复所需的部分数据处理设备，或灾难发生后能在预定时间内调配所需的数据处理设备到备用场地；要求配备部分通信线路和相应的网络设备，或灾难发生后能在预定时间内调配所需的通信线路和网络设备到备用场地。

等级三：电子传输和设备支持。要求每天至少进行一次完全数据备份，备份介质场外存放，同时每天多次利用通信网络将关键数据定时批量传送至备用场地。配备灾难恢复所需的部分数据处理设备、通信线路和相应的网络设备。

等级四：电子传输及完整设备支持。在等级三的基础上，要求配置灾难恢复所需的所有数据处理设备、通行线路和相应的网络设备，并且处于就绪或运行状态。

等级五：实时数据传输及完整设备支持。除要求每天至少进行一次完全数据备份，备份介质场外存放外，还要求采用远程数据复制技术，利用通信网络将关键数据实时复制到备用场地。

等级六：数据零丢失和远程集群支持。要求实现远程实时备份，数据零丢失；备用数据处理系统具备与生产数据处理系统一致的处理能力，应用软件是“集群的”，可实时无缝切换。

中国软件评测中心信息安全专家认为，灾难恢复能力等级越高，对于信息系统的保护效果越好，但同时成本也会急剧上升。因此，需要根据成本风险平衡原则（即灾难恢复资源的成本与风险可能造成的损失之间取得平衡），确定业务系统的合理的灾难恢复能力等级。对于多个业务系统，不同业务可采用不同的灾难恢复策略。

信息系统灾难恢复能力等级与恢复时间目标（RTO）和恢复点目标（RPO）具有一定的对应关系，各行业可根据行业特点和信息技术的应用情况制定相应的灾难恢复能力等级要求和指标要求。

### 5.1.5 信息系统应急演练

信息系统突发事件的发生是非常随机的。对于一个管理健全的信息系统，突发事件的发生概率极低。但是一旦发生就是一个非常大的灾难。这时，即使是有完美的应急预案，也会因为不太熟练或手忙脚乱，而贻误时机或处理不到位而造成系统损失。因此，光有安全管理制度、应急领导组织和应急预案还是不够，还需要通过应急演练提高应急处理的响应能力、灾难恢复能力和处置能力。

应急演练一般包含如下内容。

（1）明确应急演练的指导思想。

（2）成立应急演练组织。

（3）制定应急演练方案。

#### 1. 应急演练的指导思想

信息系统灾备演练对于检验信息系统灾难恢复预案的适用性、有效性，提升灾备系统的实际恢复能力具有重要意义。因此，应急演练不能看作仅仅是演练，而应该当作一次实战。要从实战出发，认真对待。

### 2. 应急演练组织机构

应急演练的组织一般分为应急演练指挥部和应急演练工作组。

1）应急演练指挥部

应急演练指挥部的职责是：负责信息系统突发事件应急演练的指挥、组织协调和过程控制；向上级部门报告应急演练进展情况和总结报告，确保演练工作达到预期目的。

2）应急演练工作组

应急演练工作组的职责如下。

（1）负责信息系统突发事件应急演练的具体工作，对信息系统突发事件应急演练业务影响情况进行分析和评估。

（2）收集分析信息系统突发事件应急演练处置过程中的数据信息和记录。

（3）向应急指挥部报告应急演练进展情况和事态发展情况。

（4）做好后勤保障工作，提供应急演练所需人力和物力等资源保障。

（5）做好对受影响客户的解释和安抚工作。

（6）做好秩序维护、安全保障支援等工作。

（7）建立与电力、通信、公安等相关外部机构的应急演练协调机制和应急演练联动机制。

（8）其他为降低事件负面影响或损失提供的应急支持保障等。

### 3. 制定演练方案

演练方案包含如下内容。

（1）演练时间。

（2）选定演练目的。

（3）确定演练内容。

### 4. 演练准备工作

（1）组织员工学习信息安全的有关规范和本组织的信息系统突发事件应急预案。

（2）提高员工对于突发事件的应急处置意识，熟悉在突发事件中各自的职责和任务。

（3）明确责任，严格组织实施演练活动，确保演练活动顺利完成，达到预期效果。

（4）制定演练详细时间安排表。

### 5. 应急演练的实施

根据演练方案开展演练。

### 6. 总结汇报

演练结束后，要对演练进行总结，对演练中出现的问题要及时上报并进行整改。

## 5.2 数据备份、数据容错与数据容灾

信息系统是脆弱的，它的可靠性不断遭受威胁。为了保证系统的可靠性，经过长期摸索，人们总结出了 3 条途径：避错、纠错和容错。

避错是完善设计和制造，试图构造一个不会发生故障的系统。但是，这是不太现实的。任何一个系统都会有纰漏。因此，人们不得不用纠错作为避错的补充。一旦系统出现故障，可以通过检测和核实来消除，再进行系统的恢复。

容错是第三条途径。其基本思想是，即使出现错误，系统也还能执行一组规定的程序。或者说，程序不会因为系统中的故障而中断或被修改，并且故障也不引起执行结果的差错。或者简单地说，容错就是系统可以抵抗错误的能力。容灾是针对灾害而言的。灾害对系统危害比错误要大、要严重。

### 5.2.1 数据备份

数据备份是数据容灾、数据容错和数据恢复的基本保障措施。

#### 1. 数据备份的概念

为了清楚备份的概念，需要从以下几个方面理清它与另外一个意义相近的术语——复制之间的关系。

1）从词面解释看

“备份”对应的英语单词是 backup，它在英语中具有支持、后援、备用、候补、阻塞、伴奏、倒退、裱等意思。复制对应的英语单词是 copy，它在英语中具有复制、抄写、模仿等意思。

从汉语方面看，备份就是准备好一份，以便必要时应急。所以它是基于可靠或安全进行的冗余性保存。而复制是照样做一个的意思，不一定是为了可靠与安全，也许还有别的目的。

2）从形式上看

从形式上看，复制有两个特点。

（1）与源数据内容完全相同。

（2）与源数据文件格式完全相同。

而备份与之不同：

（1）不一定是全部，可能是全部，也可能是一部分，只要应急够用就行。

（2）格式不一定相同。备份根据备份软件的不同，会被打包成不同的备份文件格式，只能用备份软件恢复过来，不能直接使用。多数备份软件（比如 VERITAS NetBackup 软件）的备份格式为标准的 TAR 格式，也就是磁带的格式。

#### 2. 数据备份模式

从模式看，数据备份可以分为逻辑备份和物理备份。

1）逻辑备份

逻辑备份也称为“基于文件（file-based）的备份”，即以文件为单位进行复制备份。这种备份，使得每个单独的文件恢复比较简单。但是，一个文件往往可能由分散在磁盘上的多个数据块连接而成；文件备份需要进行文件操作，又需要对数据块进行操作。这样，对非连续存储在磁盘上的文件进行备份时，需要额外的查找操作。这些额外的查找操作会增加磁盘开销。此外，即使文件中有一点很小的改变，也要对整个文件进行一次备份。

2）物理备份

物理备份也称为“基于块（block-based）的备份”或“基于设备（device-based）的备份”。这种备份以数据件、块为单位进行备份，因此在备份过程中，花费在搜索上的开销很少，可以提高备份的性能。但是在恢复时必须收集文件和目录的信息，要知道具体的数据块是以什么方式组织到文件中的，因此恢复的效率很低。

### 3. 数据备份环境

按照环境，备份可以是冷备份，也可以是热备份。

1）冷备份

冷备份也称为离线备份或脱机备份，是数据库已经正常关闭的情况下将关键性文件复制到另外位置的备份。这样，可以很好地解决备份选择进行时并发数据更新所带来的数据不一致。其缺点是用户需要等待较长的时间。

2）热备份

热备份也称为在线备份或同步数据备份，是在数据库运行的情况下，采用归档模式（archivelog mode）备份数据的方法。采用这种备份时，用户不需等待，即在数据更新时也允许数据备份，但要采用文件的单独写/修改特权等技术措施解决数据的不一致问题。

### 4. 数据备份的策略

数据备份策略包括备份时间、备份数据种类和故障恢复方式等。下面介绍 4 种备份策略。

1）完全备份

完全备份（full backup）指定时对整个系统或用户指定的所有文件数据进行一次完全的备份。例如，星期一用一个磁盘（或光盘）对整个系统进行一次备份，星期二再用另一个磁盘（或光盘）对整个系统进行一次备份，以此类推。这种备份策略比较可靠，可以将数据恢复到任何一次备份之前，但不能保证将数据恢复到灾难之前。并且，其成本较高，需要大量存储空间。

2）增量备份

增量备份（incremental backup）只备份上次备份后作过更新的文件。例如，星期一先用一个磁盘（或光盘）进行一次完全备份，星期二则只备份自星期一备份以后新的或被修改

过的数据，星期三只备份自星期二备份以后新的或被修改过的数据，以此类推。这种备份使系统性能和容量可以得到很好的改善。但是，一旦出现数据故障时要进行数据恢复，不得不从最后一次的备份向前进行链式恢复。例如，星期四出现故障，则要先从星期一的备份数据开始，用星期二的备份恢复出星期二备份之前的数据；再以此为基础，用星期三的备份恢复出星期三备份之前的数据。若其中任何一个中间环节出现问题，都使恢复难于继续。因此，这种模式与完全备份配合使用效果较好。

3）差别备份

差别备份（differential backup）是每次只备份上次全盘备份之后更新过的数据。例如，星期一先用一个磁盘（或光盘）进行一次完全备份，以后每天只备份与星期一的备份数据不同的数据部分。这样，全部系统只需要两组磁盘或光盘（最后一次完全备份磁盘或光盘和最后一次差别备份磁盘或光盘）就可以恢复。

4）按需备份

按需备份是指在正常的备份之外，有选择地进行的额外备份操作（例如对于非常关键的数据）。按需分配可以弥补冗余管理或长期转储的日常备份的不足。

**5. 数据备份的存储**

1）直接备份

使用备份软件直接备份在磁带、磁盘、闪存、光盘等介质上。

2）网络备份

网络备份是通过网络进行数据备份。主要形式有如下几种。

（1）网络附加存储（Network Attached Storage，NAS），其结构如图 5.2 所示，它通过在网络中安装一种只负责实现文件 I/O 操作的设施，把任务优化的存储设备直接挂在网上，使数据的存储与数据的处理相分离：文件服务器只用于数据的存储，主服务器只用于数据的处理。




图 5.2　NAS 的存储结构

（2）存储局域网（Storage Area Network，SAN）：是用来连接服务器和存储装置（大容量磁盘阵列和备份磁带库等）的专用存储网络。如图 5.3 所示，SAN 的连接基于固有光纤通道和 SCSI，通过 SCSI 到光纤通道转换器和网关，一个或多个光纤通道交换机在主服务器与存储设备之间提供相互连接，形成一种特殊的高速网络。如果把 LAN 作为第一网络，则 SAN 就是第二网络，它置于 LAN 之下，但又不涉及 LAN 的具体操作。

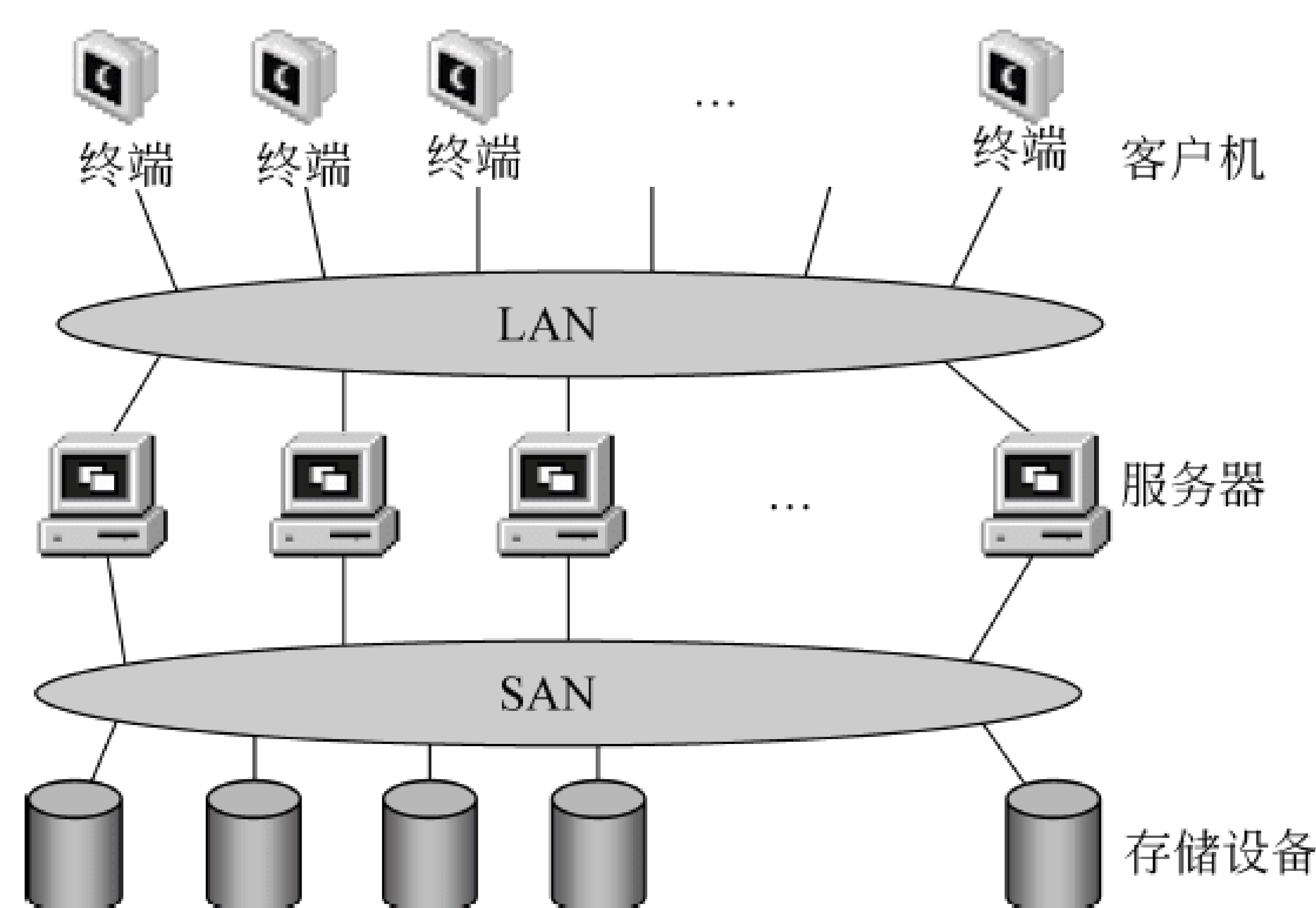



图 5.3　SAN 的结构

（3）云存储：即云计算的数据存储技术，它具有分布式、高吞吐率、大冗余和高传输率的特点。目前云计算系统中广泛使用的数据存储系统是 Google 的 GFS（Google File System，Google 文件系统，非开源）和 Hadoop 团队开发的 GFS 的开源实现 HDFS（Hadoop Distributed File System）。目前这两种技术已经成为事实标准。

GFS 是一个可扩展的分布式文件系统，用于大型的、分布式的、对大量数据进行访问的应用。一个 GFS 集群由一个主服务器（master）和大量的块服务器（chunk server）构成，并被许多客户（client）访问。主服务器存储文件系统所有的元数据（描述数据的数据），包括名字空间、访问控制信息、从文件到块的映射以及块的当前位置。它也控制系统范围的活动，如块租约（lease）管理、孤儿块的垃圾收集、块服务器间的块迁移。主服务器定期通过 HeartBeat 消息与每一个块服务器通信，给块服务器传递指令并收集它的状态。GFS 中的文件被切分为 64MB 的块并以冗余存储，每份数据在系统中保存 3 个以上备份。客户与主服务器的交换只限于对元数据的操作，所有数据方面的通信都直接和块服务器联系，这大大提高了系统的效率，防止主服务器负载过重。

### 6. 数据备份关键技术

1）镜像技术

通常，镜像（mirroring）是指一个物体对于一个镜面的重现。在计算机技术中，镜像是一种冗余的类型，是数据的另一个完全相同的副本。

（1）按照存在形式，镜像可以分为磁盘镜像和镜像站点。

磁盘镜像是在两个或多个磁盘或磁盘子系统上产生同一个数据的镜像视图的信息存储过程，即一个磁盘上的数据在另一个磁盘上存在一个完全相同的副本即为镜像。RAID 1 和 RAID 10 使用的就是镜像。常见的镜像文件格式有 iso、bin、img、tao、dao、cif、fcd。

镜像站点指某个网站存在另一个相关内容完全相同的网站。

（2）按照存在关系，镜像系统有主从之分。

在磁盘镜像系统中，要将一个磁盘指定为主磁盘镜像系统，将另一个指定为从磁盘镜像系统。

在镜像站点系统中，则将一台服务器被指定为主服务器，将另一台指定为从服务器。

客户只能对主服务器上的镜像的卷进行读写，即只有主服务器通过网络向用户提供服务，从服务器上相应的卷被锁定以防对数据的存取。主/从服务器分别通过心跳监测线路互相监测对方的运行状态，当主服务器因故障停机时，从服务器将在很短的时间内接管主服务器的应用。

（3）按主从磁盘镜像存储系统所处的位置，磁盘镜像可分为本地镜像和远程镜像。

本地镜像是在本机的磁盘中划分主镜像区和从镜像区。远程镜像又称为远程复制，是通过高速光纤通道线路和磁盘控制技术将镜像磁盘延伸到远离本地机的地方，镜像磁盘数据与主磁盘数据完全一致。

（4）按请求镜像的主机是否需要远程镜像站点的确认信息，远程镜像又可分为同步远程镜像和异步远程镜像。

同步远程镜像（同步复制技术）是指通过远程镜像软件，将本地数据以完全同步的方式复制到异地，每一本地的 I/O 事务均需等待远程复制的完成确认信息，方予以释放。同步镜像使复制总能与本地机要求复制的内容相匹配。当主站点出现故障时，用户的应用程序切换到备份的替代站点后，被镜像的远程副本可以保证业务继续执行而没有数据的丢失。但它存在往返传播造成延时较长的缺点，只限于在相对较近的距离上应用。

异步远程镜像（异步复制技术）保证在更新远程存储视图前完成向本地存储系统的基本操作，而由本地存储系统提供给请求镜像主机的 I/O 操作完成确认信息。远程的数据复制是以后台同步的方式进行的，这使本地系统性能受到的影响很小，传输距离长（可达 1000km 以上），对网络带宽要求低。但是，许多远程的从属存储子系统的写没有得到确认，当某种因素造成数据传输失败，可能出现数据一致性问题。为了解决这个问题，目前大多采用延迟复制的技术（本地数据复制均在后台日志区进行），即在确保本地数据完好无损后进行远程数据更新。

2）存储快照技术

快照（snapshot）是通过软件对磁盘子系统中的数据快速扫描，为要备份数据在某个时间点上建立的映像，这个映像由快照逻辑单元号（LUN）和快照高速缓存组成。快照 LUN 是一组指针，它指向快照高速缓存和磁盘子系统中不变的数据块（在备份过程中）。在快速扫描时，把备份过程中即将要修改的数据块同时快速复制到快照高速缓存中。

快照能够进行在线数据备份与恢复。当存储设备发生应用故障或者文件损坏时可以进行快速的数据恢复，将数据恢复至某个可用的时间点的状态。在正常业务进行的同时，利用快照 LUN 实现对原数据的一个完全的备份。它可使用户在正常业务不受影响的情况下（主要指容灾备份系统），实时提取当前在线业务数据。

快照的另一个作用是为存储用户提供另外一个数据访问通道，当原数据进行在线应用处理时，用户可以访问快照数据，还可以利用快照进行测试等工作。由于其“备份窗口”接近于零，可大大增加系统业务的连续性，为实现系统真正的 7×24 小时运转提供了保证。所有存储系统，不论高中低端，只要应用于在线系统，那么快照就成为一个不可或缺的功能。

远程镜像技术往往同快照技术结合起来实现远程备份，即通过镜像把数据备份到远程

存储系统中，再用快照技术把远程存储系统中的信息备份到远程的磁带库、光盘库中。

3）互连技术

早期的主数据中心和备援数据中心之间的数据备份，主要是基于 SAN（存储区域网络）的远程复制（镜像），即通过光纤通道（FC），把两个 SAN 连接起来，进行远程镜像（复制）。当灾难发生时，由备援数据中心替代主数据中心保证系统工作的连续性。但是由于这种远程容灾备份方式存在实现成本高、设备的互操作性差、跨越的地理距离短（10km）等因素，阻碍了它的进一步推广和应用。

目前出现了多种基于 IP 的 SAN 的远程数据容灾备份技术。它们是利用基于 IP 的 SAN 的互连协议，将主数据中心 SAN 中的信息通过现有的 TCP/IP 网络远程复制到备援中心 SAN 中。当备援中心存储的数据量过大时，可利用快照技术将其备份到磁带库或光盘库中。这种基于 IP 的 SAN 的远程容灾备份，可以跨越 LAN、MAN 和 WAN，成本低，可扩展性好，具有广阔的发展前景。基于 IP 的互连协议包括 FCIP、iFCP、Infiniband 和 iSCSI 等。

4）虚拟存储

虚拟存储就是把多个存储介质模块（如硬盘、RAID）通过一定的手段集中在一个存储池（storage pool）中进行统一管理。这样，从主机的角度看到的就不是多个硬盘，而是一个分区或者卷。从拓扑结构来讲，虚拟存储主要有两种方式：对称式虚拟存储和非对称式虚拟存储。对称虚拟存储有如下优点。

（1）采用大容量高速缓存，可以显著提高数据传输速度。

（2）采用多端口并行技术，可以消除 I/O 瓶颈。

（3）其逻辑存储单元提供了高速的磁盘访问速度。

（4）成对的存储池具有很好的容错性能。

### 5.2.2　数据容错技术

容错（Fault Tolerant，FT）就是当由于种种原因在系统中出现了数据、文件损坏或丢失时，系统能够自动将这些损坏或丢失的文件和数据恢复到发生事故以前的状态，使系统能够连续正常运行的技术。

目前，广泛采用的数据容错技术有以下几种。

#### 1. 双重文件分配表和目录表技术

硬盘上的文件分配表和目录表存放着文件在硬盘上的位置和文件大小等信息，如果它们出现故障，数据就会丢失或误存到其他文件中。通过提供两份同样的文件分配表和目录表，把它们存放在不同的位置，一旦某份出现故障，系统将做出提示，从而达到容错的目的。

#### 2. 快速磁盘检修技术

这种方法是在把数据写入硬盘后，马上从硬盘中把刚写入的数据读出来与内存中的原始数据进行比较。如果出现错误，则利用在硬盘内开设的一个被称为“热定位重定区”的区，将硬盘坏区记录下来，并将已确定的在坏区中的数据用原始数据写入热定位重定区上。

**3．磁盘镜像技术**

磁盘镜像是在同一存储通道上装有成对的两个磁盘驱动器，分别驱动原盘和副盘，两个盘串行交替工作，当原盘发生故障时，副盘仍旧正常工作，从而保证了数据的正确性。

**4．双工磁盘技术**

它是在网络系统上建立起两套同样的且同步工作的文件服务器，如果其中一个出现故障，另一个将立即自动投入系统，接替发生故障的文件服务器的全部工作。

**5．事务跟踪系统**

网络操作系统具有完备的事务跟踪系统，这是针对数据库和多用户软件的需要而设计的，用以保证数据库和多用户应用软件在全部处理工作还没有结束时，或者在工作站或服务器发生突然损坏的情况下，能够保持数据的一致。其工作方式是：对指定的事务（操作）要么一次完成，要么什么操作也不进行。

**6．负载均衡**

负载均衡（load balance）就是将一个任务分解成多个子任务，分配给不同的服务器执行，通过减少每个部件的工作量，增加系统的稳定性。通常，负载均衡会根据网络的不同层次（网络七层）来划分。其中，第二层的负载均衡指将多条物理链路当作一条单一的聚合逻辑链路使用，这就是链路聚合（trunking）技术，它不是一种独立的设备，而是交换机等网络设备的常用技术。现代负载均衡技术通常操作于网络的第四层或第七层，它完全脱离交换机、服务器而成为独立的技术设备。

**7．LOCKSTEP 技术**

LOCKSTEP 技术使用相同的、冗余的硬件组件在同一时间内处理相同的指令。它可以保持多个 CPU、内存精确的同步，在相同时钟周期内执行相同的指令；保证能够发现任何错误，即使短暂的错误，系统也能在不间断处理和不损失数据的情况下恢复正常运行。

**8．安全故障（FAILSAFE）软件**

FAILSAFE 可以管理和诊断特征，捕获、分析和通报服务器的软件问题，从而允许个人在软件发生错误之前去纠正错误。FAILSAFE 软件通过下列功能来增强 Windows 环境中的可靠性。

（1）保护短暂的硬件故障。

（2）通过增强的驱动程序预防软件失效。

（3）软件问题的捕获、分析及修正。

（4）内存数据的连续性维持。

（5）丰富的纠错功能可以解决各种不同的错误。

（6）自动重启功能，能够将宕机前 CPU 与内存数据即时保存下来。

FAILSAFE 软件在 Windows 2000/2003 环境下采用热插拔、内存镜像、负载均衡、多点

终止失效、多通道 I/O 等方式，大大增强了系统连续运行的稳定性。

#### 9. 服务器容错技术

服务器容错技术的出现极大地降低了企业业务在各种不可预料灾难发生时的损失，保证业务系统的 7×24 小时不间断运转。常用的服务器容错技术有下列一些。

（1）双机热备（hot standby）：通过系统冗余的方法解决计算机应用系统的可靠性问题，并具有安装维护简单、稳定可靠、监测直观等优点。

（2）服务器集群（cluster）：服务器集群是通过将多台服务器互连在一起而形成的，以松散的成对配置共享资源，具有一定的自我修正能力，可以保证系统 7×24 小时不间断运行，把非计划和计划的停机时间降到最低。在确保高可用性方面，服务器集群堪称最具价值的系统级技术之一。

（3）SAN：允许服务器在共享存储装置的同时仍能高速传送数据，具有带宽高、可用性高、容错能力强的优点，而且它可以轻松升级，容易管理，有助于改善整个系统的总体成本状况。

### 5.2.3 数据容灾系统

真正的数据容灾就是要能在灾难发生时全面、及时地恢复整个系统。在系统遭受灾害时，使系统还能工作或尽快恢复工作的最基础的工作是数据备份。对于一个容灾系统，如果没有备份的数据，任何容灾方案都没有现实意义。

#### 1. 衡量容灾的两个技术指标

从技术上看，衡量容灾系统有两个主要指标——RPO 和 RTO。

数据恢复点目标（Recovery Point Object，RPO）主要指的是业务系统所能容忍的数据丢失量，代表了当灾难发生时允许丢失的数据量。

恢复时间目标（Recovery Time Object，RTO）主要指的是所能容忍的业务停止服务的最长时间，代表了系统恢复的时间，也就是从灾难发生到业务系统恢复服务功能所需要的最短时间周期。

所以，RPO 针对的是数据丢失，而 RTO 针对的是服务停止，两者没有必然的关联性。RTO 和 RPO 的确定必须在进行风险分析和业务影响分析后根据不同的业务需求确定。对于不同企业的同一种业务，RTO 和 RPO 的需求也会有所不同。

#### 2. 容灾必须满足的 3R

真正的容灾必须满足 3R。

（1）Redundance，即系统中的部件、数据都具有“冗余性”，即一个系统发生故障，另一个系统能够保持数据传送的顺畅。

（2）Remote，即具有“长距离性”。因为灾害总是在一定范围内发生，因而充分长的距离才能够保证数据不会被一个灾害全部破坏。

（3）Replication，容灾系统要追求全方位的数据“备份”，也称为容灾性。

### 3. 确定容灾备份技术方案的因素

根据国际标准 SHARE 78 的定义，确定灾难备份技术方案应主要考虑如下 8 个方面。

（1）备份/恢复的范围。

（2）灾难恢复计划的状态。

（3）应用站点与灾难备份站点之间的距离。

（4）应用站点与灾难备份站点之间是如何相互连接的。

（5）数据是怎样在两个站点之间传送的。

（6）允许有多少数据被丢失。

（7）怎样保证更新的数据在灾难备份站点被更新。

（8）灾难备份站点可以开始灾难备份工作的能力。

### 4. 七级容灾系统等级

按照以上考虑因素，国际标准 SHARE 78 将容灾系统定义成如下 7 个层次。

1）0 级：无异地备份

在这种容灾方案中，最常用的是备份管理软件加上磁带机，可以手工加载磁带机或自动加载磁带机。其特点是用户投资较少，技术实现简单。缺点是一旦本地发生毁灭性灾难，将丢失全部的本地备份数据，业务无法恢复，所以不具备真正的灾难恢复能力。

2）1 级：实现异地备份

这种方案是将关键数据备份到本地磁带介质上，然后送往异地保存，但异地没有可用的备份中心、备份数据处理系统和备份网络通信系统，未制订灾难恢复计划。灾难发生后，使用新的主机，利用异地数据备份介质（磁带）将数据恢复起来。这种方案成本较低，但难以管理，即很难知道什么数据在什么地方，恢复时间长短依赖于何时硬件平台能够被提供和准备好。目前，这一等级方案在许多中小网站和中小企业用户中采用较多。

3）2 级：热备份站点备份

这种方案是将关键数据进行备份并存放到异地，制订有相应灾难恢复计划，具有热备份能力的站点灾难恢复。一旦发生灾难，利用热备份主机系统将数据恢复。由于有了热备中心，用户投资会增加，相应的管理人员要增加。技术实现简单，利用异地的热备份系统，可以在本地发生毁灭性灾难后，快速进行业务恢复。但这种容灾方案由于备份介质是采用交通运输方式送往异地，异地热备中心保存的数据是上一次备份的数据，可能会有几天甚至几周的数据丢失。这对于关键数据的容灾是不能容忍的。

4）3 级：在线数据恢复

这种方案制订有相应灾难恢复计划，有备份中心，配备有部分数据处理系统及网络通信系统，并通过网络将关键数据进行备份并存放至异地，一旦灾难发生，需要的关键数据通过网络可迅速恢复，通过网络切换，关键应用恢复时间可降低到一天或小时级。但由于备份站点要保持持续运行，对网络的要求较高，因此成本相应有所增加。

5）4 级：定时数据备份

这一方案是在第 3 级容灾方案的基础上，利用备份管理软件自动通过通信网络将部分关键数据定时备份至异地，并制订相应的灾难恢复计划；异地热备中心保存的数据是定时备份的数据，根据备份策略的不同，数据的丢失与恢复时间达到天或小时级。一旦灾难发生，利用热备中心已有资源及异地备份数据恢复关键业务系统运行。但投入成本也会增加。

6）5 级：实时数据备份

这一容灾方案在前面几个级别的基础上使用了硬件的镜像技术和软件的数据复制技术，也就是说，可以实现在应用站点与备份站点的数据都被更新。数据在两个站点之间相互镜像，由远程异步提交来同步，因为关键应用使用了双重在线存储，所以在灾难发生时，仅仅很小部分的数据被丢失，恢复的时间被降低到了分钟级或秒级。由于对存储系统和数据复制软件的要求较高，所需成本也大大增加。

7）6 级：零数据丢失

第 6 级容灾方案是灾难恢复中最昂贵的方式，也是速度最快的恢复方式，它是灾难恢复的最高级别，利用专用的存储网络将关键数据同步镜像至备份中心，数据不仅在本地进行确认，而且需要在异地（备份）进行确认。因为，数据是镜像地写到两个站点，所以灾难发生时异地容灾系统保留了全部的数据，实现零数据丢失。

这 7 个层次对应的容灾方案在功能、适用范围等方面都有所不同，所以用户在选型时应分清层次。

**5. 简化的容灾系统等级**

由于上述 7 层比较复杂，人们常简化为如下 4 个等级。

1）第 0 级：没有备援中心

这一级容灾备份实际上没有灾难恢复能力，它只在本地进行数据备份，并且被备份的数据只在本地保存，没有送往异地。

2）第 1 级：本地磁带备份，异地保存

在本地将关键数据备份，然后送到异地保存。灾难发生后，按预定数据恢复程序恢复系统和数据。这种方案成本低、易于配置。但当数据量增大时，存在存储介质难管理的问题，并且当灾难发生时存在大量数据难以及时恢复的问题。为了解决此问题，灾难发生时，先恢复关键数据，后恢复非关键数据。

3）第 2 级：热备份站点备份

在异地建立一个热备份点，通过网络进行数据备份。也就是通过网络以同步或异步方式，把主站点的数据备份到备份站点，备份站点一般只备份数据，不承担业务。当出现灾难时，备份站点接替主站点的业务，从而维护业务运行的连续性。

4）第 3 级：活动备援中心

在相隔较远的地方分别建立两个数据中心，它们都处于工作状态，并进行相互数据备

份。当某个数据中心发生灾难时，另一个数据中心接替其工作任务。这种级别的备份根据实际要求和投入资金的多少，又可分为两种：①两个数据中心只限于关键数据的相互备份；②两个数据中心互为镜像，即零数据丢失等。零数据丢失是目前要求最高的一种容灾备份方式，它要求不管什么灾难发生，系统都能保证数据的安全。所以，它需要配置复杂的管理软件和专用的硬件设备，需要的投资相对而言是最大的，但恢复速度也是最快的。

#### 6. 容灾方案

目前有很多种容灾技术，分类也比较复杂。但总体上可以区分为离线式容灾（冷容灾）和在线容灾（热容灾）两种类型。

1）离线式容灾（冷容灾）

离线式容灾主要依靠备份技术来实现。其重要步骤是：将数据通过备份系统备份到磁带上面，然后将磁带运送到异地保存管理。这种方式主要由备份软件来实现备份和磁带的管理，除了磁带的运送和存放外，其他步骤可实现自动化管理。整个方案的部署和管理比较简单，相应的投资也较少。但缺点也比较明显：由于采用磁带存放数据，所以数据恢复较慢，而且备份窗口内的数据都会丢失，实时性比较差。对于资金受限、对数据恢复的 RTO 和 RPO 要求较低的用户可以选择这种方式。

2）在线式容灾（热容灾）

在线式容灾要求用户工作中心和灾备中心同时工作，用户工作中心和灾备中心之间有传输链路连接。数据自用户工作中心实时复制传送到灾备中心。在此基础上，可以在应用层进行集群管理，当用户工作中心遭受灾难、出现故障时，可由灾备中心自动接管并继续提供服务。应用层的管理一般由专门的软件来实现，可以代替管理员实现自动管理。由上面分析可见，实现在线式容灾的关键是数据的复制。

#### 7. 在线式容灾的数据复制方式

数据复制的技术有很多，从实现复制功能的设备分布上可大体分为 3 层：服务器层、存储交换机层和存储层。

1）服务器层

在用户工作中心和灾备中心的服务器上安装专用的数据复制软件，以实现远程复制功能。两中心间必须有网络连接作为数据通道。可以在服务器层增加应用远程切换功能软件，从而构成完整的应用级容灾方案。这种数据复制方式相对投入较少，主要是软件的采购成本。另外，其兼容性较好，可以兼容不同品牌的服务器和存储设备，较适合硬件组成复杂的用户。但这种方式要在服务器上运行软件，不可避免地对服务器性能产生影响。

2）存储交换机层

存储交换机技术的发展使交换机可以实现更多的功能，很多原来由服务器和存储设备实现的功能现在也可在交换机层实现，例如存储虚拟化。由于交换机可以管理和复制的数据是存放在存储层的，因此用户需要将数据都存储在交换机所连接的存储设备中，这样就

可以实现交换机对数据的管理和复制。目前采用这种方案的用户比较少。

3）存储层

远程数据复制功能几乎是现有中高端产品的必备功能。要实现数据的复制，需要在用户工作中心和灾备中心都部署一套这样的存储系统，数据复制功能由存储系统实现。距离比较近（几十千米之内）时，之间的链路可由两中心的存储交换机通过光纤直接连接；距离在 200km 内时，可通过增加 DWDM 等设备直接进行光纤连接；距离超过 200km，则可增加存储路由器进行协议转换，途经 WAN 或 Internet 实现连接。因此，从理论上可实现无限制连接。在存储层实现数据复制功能是很成熟的技术，而且对应用服务器的性能基本没有影响。在应用层增加远程集群软件后就可以实现自动灾难切换的整体容灾解决方案。这种容灾方案稳定性高，对服务器性能基本无影响，是目前容灾方案的主流选择。

#### 8. 灾难检测技术

对于一个容灾系统来讲，在灾难发生时，尽早地发现生产系统端的灾难，尽快地恢复生产系统的正常运行或者尽快地将业务迁移到备用系统上，都可以将灾难造成的损失降低到最低。除了依靠人力来对灾难进行确定之外，对于系统意外停机等灾难还需要容灾系统能够自动地检测灾难的发生，目前容灾系统的检测技术一般采用心跳技术。

心跳技术的一个实现是：生产系统在空闲时每隔一段时间向外广播一下自身的状态；检测系统在收到这些“心跳信号”之后，便认为生产系统是正常的。若在给定的一段时间内没有收到“心跳信号”，检测系统便认为生产系统出现了非正常的灾难。心跳技术的另外一个实现是：每隔一段时间，检测系统就对生产系统进行一次检测，如果在给定的时间内，被检测的系统没有响应，则认为被检测的系统出现了非正常的灾难。心跳技术中的关键点是心跳检测的时间和时间间隔周期。如果间隔周期短，会给系统带来很大的开销。如果间隔周期长，则无法及时地发现故障。

#### 9. 系统迁移技术

灾难发生后，为了保持生产系统的业务连续性，需要实现系统的透明性迁移，利用备用系统透明地代替生产系统进行运作。一般对实时性要求不高的容灾系统，例如 Web 服务、邮件服务器等，可以通过修改 DNS 或者 IP 来实现，对实时性要求高的容灾系统，则需要将生产系统的应用透明地迁移到备用系统上。目前基于本地机群的进程迁移的算法可以应用在远程容灾系统中，但是需要对迁移算法进行改进，使之适应复杂的网络环境。

## 5.3　数字证据获取

现在，信息系统的攻击和对抗已经不仅仅是技术领域和管理领域的问题了。许多问题已经涉讼，成为法学案件。随着数字犯罪案件的增多，数字证据的获取已经成为信息技术和法学家们共同关注的热点。

数字证据也称为计算机证据。对于它的研究最早是从应急响应的角度开始的，目的是为了搜集攻击者的有关信息。直到 2001 年人们才转移到从司法的角度来看待它，关于它的

研究，才从纯技术领域转向技术与法学的结合上。

### 5.3.1 数字证据的特点与数字取证的基本原则

#### 1. 数字证据的特点

数字证据就在计算机或计算机系统运行过程中产生的、以其记录的内容来证明案件事实的电磁记录。与其他证据相比，它有如下一些特点。

1）依附性和多样性

电磁证据依附在不同介质上。这就带来两个方面的特点：一是数字证据不会像传统的证据那样可以独立存在；二是不同的介质使同样的信息表现出不同的形态，如在导体中是以电流或电压表现的数字脉冲，在显示器上是文字或图形，在磁盘中是磁核的排列形式，在光缆中是光波等。

2）可伪性和弱证明性

数字证据的非实物性，使得其窃取、修改甚至销毁都比较容易。例如，黑客在入侵之后，可以对现场进行一些灭迹、制造假象等工作，给证据的认定带来困难，直接减低了证明力度，增加了跟踪和侦查的难度。

3）数据的挥发性

计算机系统中所处理的数据有一些是动态的。这些动态数据对于发现犯罪的蛛丝马迹非常有用。但是它们却有一定的时间效应，即有些数据会因失效或消失而挥发。在收集数字证据时必须充分考虑数据的挥发性。表 5.1 描述了数字证据数据的挥发性。

**表 5.1 数字证据数据的挥发性**

| 数 据 | 硬件或位置 | 存 活 时 间 |
|---|---|---|
| CPU | 高速缓冲器，管道 | 几个时钟周期 |
| 系统 | RAM | 关机前 |
| 内核表 | 进程中 | 关机前 |
| 固定介质 | Swap/tmp | 直至被覆盖或被抹掉 |
| 可移动介质 | CD-ROM，Floppy，HDO | 直至被覆盖或被抹掉 |
| 打印输出 | 被打印输出 | 直至被毁坏 |

#### 2. 数字取证的基本原则

实施数字取证应当遵循如下原则：符合程序，共同监督，保护隐私，影响最小，连续完全，原汁原味。下面分别予以说明。

1）符合程序

取证应当首先启动法律程序，要在法律规定的范围内展开工作，否则会陷入被动。

2）共同监督

由原告委派的专家所进行的整个检查、取证过程必须受到由其他方委派的专家的监督。

3）保护隐私

在取证过程中，要尊重任何关于客户代理人的隐私。一旦获取了一些关于公司或个人的隐私，绝不能泄露。

4）影响最小

（1）如果取证要求必须运行某些业务程序，应当使运行时间尽量短。
（2）必须保证取证不给系统带来副作用，如引进病毒等。

5）连续完全

必须保证证据的连续性（chain of custody），即在将证据提交法庭前要一直跟踪证据，要向法庭说明在这段时间内证据有无变化。此外，要向法庭说明该证据的完全性。

6）原汁原味

（1）必须保证提取出来的证据不受电磁或机械的损害。
（2）必须保证收集的证据不被取证程序破坏。

### 5.3.2 数字取证的一般步骤

数字取证过程一般可以按如下步骤进行。

#### 1. 保护现场

（1）在取证过程中，保护目标系统，避免发生任何改变、损害。
（2）保护证据的完整性，防止证据信息的丢失和破坏。
（3）防止病毒感染。

#### 2. 证据发现

证据发现首先要识别可获取证据的信息类型。按照证据信息变化的特点，可以将取证分为两大类。

（1）来源取证。即确定犯罪嫌疑人或者证据的来源。例如，在网络犯罪侦查中，为了确定犯罪嫌疑人，可能需要找到犯罪嫌疑人犯罪时使用的机器的 IP 地址，则寻找 IP 地址便是来源取证。这类取证中，主要有 IP 地址取证、MAC 地址取证、电子邮件取证、软件账号取证等。

（2）事实取证。即不是为了查明犯罪嫌疑人，而是取得与证明案件相关事实的证据，例如犯罪嫌疑人的犯罪事实证据。在事实取证中常见的取证方法有文件内容调查、使用痕迹调查、软件功能分析、软件相似性分析、日志文件分析、网络状态分析、网络数据包分析等。

下面是可以作为证据或可以提供相关信息的信息源。

（1）日志，如操作系统日志等。
（2）文件，如可以进行的文件搜索有以下几种。
① 搜索目标系统中的所有文件（包括现存的正常文件、已经被删除但仍存在于磁盘上

还没有被覆盖的文件、隐藏文件、受密码保护的和加密文件）。

② 尽量恢复所发现的文件。

③ 在法律允许的情况下，访问被保护或加密的文件。

④ 分析磁盘特殊区域（未分配区域、文件栈区等）。

（3）系统进程，如进程名、进程访问文件等。

（4）用户，特别是在线用户的服务时间、使用方式等。

（5）系统状态，如系统开放的服务、网络运行的状态等。

（6）通信连接记录，如网络路由器的运行日志等。

（7）存储介质，如磁盘、光盘、闪存等。

在证据发现阶段可以使用的技术有 IDS、蜜罐技术、网络线索自动识别技术和溯源技术等。同时还可以使用一些相关的工具。表 5.2 为一些常用实时取证类工具。

**表 5.2　一些常用实时取证类工具**

| 工具名称 | 用途描述 |
|---|---|
| EnCase | 一个集成的、基于 Windows 的取证应用程序，功能包括数据浏览、搜索、磁盘浏览、数据预览、建立案例、建立证据文件、保存案例等 |
| X-Ways Capture | 一套专业的计算机取证工具，用于在证据采集阶段获取正在运行的 Windows 和 Linux 系统下的硬盘、文件和 RAM 数据 |
| FTK | 美国警方标准配备、全球警方使用量较多的电子证据分析软件 |
| 效率源 DataCompass | 第四代专业数据恢复工具：用于硬盘数据和 U 盘数据恢复 |
| CRCMD5 | 一个可以验证一个或多个文件内容的 CRC 工具 |
| DiskSig | 一个 CRC 程序，用于验证映像备份的精确性 |
| Filter_we | 一种用于周围环境数据的智能模糊逻辑过滤器 |
| GetSlack | 一种周围环境数据收集工具，用于捕获未分配的数据 |
| GetTime | 一种周围环境数据收集工具，用于捕获分散的文件 |
| Net Threat Analyzer | 网络取证分析软件，用于识别公司互联网络账号滥用 |
| Seized | 一种用于对证据计算机上锁及保护的程序 |
| ShowFL | 用于分析文件输出清单的程序 |
| TextSearch Plus | 用来定位文本或图形文件中的字符串的工具 |

### 3. 证据固定

针对数字证据的挥发性，数字证据的固定非常重要。

### 4. 证据提取

证据提取主要是提取特征。提取方法如下。

（1）过滤和挖掘。

（2）解码：对软件或数据碎片进行残缺分析、上下文分析，恢复原来的面貌。

**5. 证据分析**

证据分析的目的大致有以下几个方面。

（1）犯罪行为重构。

（2）嫌疑人画像。

（3）确定犯罪动机。

（4）受害程度行为分析等。

**6. 提交证据**

向律师、管理者或法庭提交证据。这时要注意使用规定的法律文书格式和术语。

### 5.3.3 数字取证的基本技术和工具

在数字取证过程中，可以使用相关的技术和工具。现在已经开发出了这样一些工具。表 5.3 为可以提供计算机及网络攻击取证的一些网站。

**表 5.3 提供计算机及网络攻击取证工具的一些网站**

| 资 源 类 型 | 网 络 地 址 |
|---|---|
| TCT 取证软件包 | http://www. fish.com/forensics/ |
| Encase | http://www.encase.com/ |
| 计算机取证分析 | http://www. porcupine.org/forensics/ |
| Computer Forensics Tool Testing（CFTT） | http://www.cftt.nist.gov/ |
| 文件及介质取证工具箱 Sleuth Kit | http://www. sleuthkit.org/sleuthkit/index.php |
| 开放源代码数字取证 | http://www.opensourceforensics.org/ |

下面重点介绍利用 IDS 和蜜罐取证的方法。

**1. 利用 IDS 取证**

把 IDS 与取证工具结合，往往能对网络攻击进行取证并得到响应。

1）确认攻击

确认攻击是响应的第一步，其主要方法是查找攻击留下的痕迹。检查的主要内容如下。

（1）寻找嗅探器（如 Sniffer）。

（2）寻找远程控制程序（如 netbus、back orifice）。

（3）寻找黑客可能利用的文件共享或通信程序（如 eggdrop、irc）。

（4）寻找特权程序（如 find/-perm-4000-print）。

（5）寻找未授权的服务（如 netstat -a、check inetd.conf）。

（6）寻找异常文件（考虑系统磁盘大小）。

（7）检查文件系统的变动。

（8）检查口令文件的变动并寻找新用户。

（9）检测 cron。

（10）核对系统和网络配置（特别要注意过滤规则）。
（11）检查所有主机（特别是服务器）。

2）取证过程

（1）决定取证的目的。
① 观察研究攻击者。
② 跟踪并驱赶攻击者。
③ 捕俘攻击者。
④ 准备起诉攻击者。
（2）启动必要的法律程序。
（3）对系统进行完全备份，包括：
① 用 tcpdump 作完全的分组日志。
② 有关协议分组的来龙去脉。
③ 一些会话（如 Telnet、rlogin、IRC、FTP 等）的可能内容。
（4）根据情况有选择地关闭计算机系统。
① 不彻底关闭系统（否则造成信息改变，证据破坏）。
② 不断开网络。
③ 将系统备份转移到单用户模式下制作和验证备份。
④ 考虑制作磁盘镜像。
⑤ 同步磁盘，暂停系统。
（5）调查攻击者来源。
① 利用 tcpdump/who/syslog。
② 运行 finger 对抗远程系统。
③ 寻找攻击者可能利用的账号。

**2. 利用蜜罐取证**

利用蜜罐进行取证分析的一些原则和步骤如下。
（1）获取入侵者信息。
（2）获取关于攻击的信息。
① 攻击的手段、日期和时间。
② 入侵者添加了一些什么文件？
③ 是否安装了嗅探器或密码？若有，在何处？
④ 是否安装有 rootkit 或木马程序？若有，传播途径是什么？
⋮
（3）建立事件的时间序列。
（4）事故费用分析。
（5）向管理层、媒体以及法庭提交相应的报告。

### 5.3.4 数字证据的法律问题

数字证据学是涉及信息技术和法学两个领域的交叉学科。下面讨论它在法学方面的一些问题。

#### 1. 法律对证据的基本要求

法律对作为定案依据的证据，有真实性、合法性和关联性 3 方面要求。

一般而言，关联性主要指证据与案件争议和理由的联系程度，这属于法官裁判的范围。合法性主要包括证据的形式、收集手段、是否侵犯他人权益、取证工具是否合法等。这在后面要进行有关的讨论。

关于证据的真实性，民事诉讼法和相关司法解释都要求提供“原件”（书面文件）。因为这种看得见、摸得着的东西才能给人充分的真实感和唯一性，才能防止被篡改和冒认。而数字证据的真实性的确认一直是人们最关注的问题。目前，人们想用数字签名的方法来解决这一法律难题。通过数字签名，可以证明签发数字证据的人是谁，也可以证明数字证据是否被篡改过。

#### 2. 关于数字证据的证明力

证据的证明力是指证据对证明案件事实所具有的效力，即该证据是否能够直接证明案件事实还需要其他证据配合综合认定。《中华人民共和国刑事诉讼法》（2013 版）第五章第四十八条规定，可以用于证明案件事实的材料都是证据，包括：

（1）物证。

（2）书证。

（3）证人证言。

（4）被害人陈述。

（5）犯罪嫌疑人、被告人供述和辩解。

（6）鉴定意见。

（7）勘验、检查、辨认、侦查实验等笔录。

（8）视听资料、电子数据。

#### 3. 关于数字取证工具的法律效力

数字证据的合法性涉及数字取证工具的法律效力，即法庭是否认可。每种工具都是一个程序。按照 Daubert 测试，可以从下面 4 个方面进行讨论。

1）可测试性

测试的目的是为了确定一个程序是否可以被测试并确定它所提供的结果的准确性。对一个工具必须执行两类测试。

（1）漏判测试：确认取证工具是否可以在输入输出端提取所有可以得到的数据。

（2）误判测试：确认取证工具在输入输出端没有引入新的数据。

2）错误率

测试用于识别在数字取证工具中是否存在已知的错误。在数字取证工具中，可能存在两类错误。

（1）工具执行错误：源于代码中的漏洞的错误。

（2）提取错误：源自算法的错误。

3）公开性

公开性指工具在公开地方有证明并经过对等部门复查。这是证据得以承认的主要条件。

4）可接受性

可接受性指工具能够被广泛接受。

### 5.3.5 日志

#### 1. 日志及其用途

日志（log）是系统所指定对象的某些操作和其操作结果按时间有序的集合，是记录信息系统安全状态和问题的原始数据。通常情况下，系统日志是用户可以直接阅读的文本文件。每个日志文件由日志记录组成，每条日志记录描述了一次单独的系统事件。典型的日志内容有以下几种。

（1）事件的性质：数据的输入和输出，文件的更新（改变或修改），系统的用途或期望。

（2）全部相关标识：人、设备和程序。

（3）有关事件的信息：日期和时间，成功或失败，涉及因素的授权状态，转换次数，系统响应，项目更新地址，建立、更新或删除信息的内容，使用的程序，兼容结果和参数检测，侵权步骤等。对大量生成的日志要适当考虑数据的保存期限。

日志文件中的记录可提供以下用途。

（1）监控系统资源；审计用户行为；对可疑行为进行告警。

（2）确定入侵行为的范围；为恢复系统提供帮助；生成调查报告。

（3）为打击计算机犯罪提供证据来源。

#### 2. 日志的特点

（1）数据量大。

通常对外服务产生的日志文件，如 Web 服务日志、防火墙、入侵检测系统日志和数据库日志以及各类服务器日志等都很大，一个日志文件一天产生的容量少则几十 MB、几百 MB，多则有几个 GB、几十个 GB。

（2）日志仅反映本系统的某些特定事件的操作情况。

一个系统的日志是对本系统涉及的运行状况的信息按时间顺序作一简单的记录，仅反映本系统的某些特定事件的操作情况，并不完全反映某一用户的整个活动情况。一个用户在网络活动的过程中会在很多的系统日志中留下痕迹，如防火墙 IDS 日志、操作系统日志等，这些不同的日志之间存在某种必然的联系来反映用户的活动情况。只有将多个系统的日志

结合起来分析，才能准确反映用户的活动情况。

（3）产生系统日志的软件通常为应用系统。

产生系统日志的软件通常为应用系统而不是作为操作系统的子系统运行，所产生的日志记录容易遭到恶意的破坏或修改。系统日志通常存储在系统未经保护的目录中，并以文本方式存储，未经加密和校验处理，没有提供防止恶意篡改的有效保护机制。因此，日志文件并不一定是可靠的，入侵者可能会篡改日志文件，从而不能被视为有效的证据。由于日志是直接反映入侵者痕迹的，在计算机取证中扮演着重要的角色，入侵者获取系统权限窃取机密信息或破坏重要数据后往往会修改或删除与其相关的日志信息，甚至根据系统的漏洞伪造日志以迷惑系统管理员和审计。

（4）日志记录不会长期保存。

系统日志是根据日志文件占用磁盘空间的大小来自动删除旧的日志记录，例如，如果设置日志文件占用磁盘空间为 2MB，那么，当日志文件达到 2MB 大小时，新的日志记录会自动替换最旧的日志记录。

### 3．Windows 日志

1）日志类型及其组成

以 Windows 2000/XP 为例，日志文件通常有应用程序日志、安全日志、系统日志、DNS 服务器日志、FTP 日志和 WWW 日志等。默认文件大小为 512KB，但管理员可以改变这个默认值。

（1）应用程序日志：记录由应用程序产生的事件。例如，某个数据库程序可能设定为每次成功完成备份操作后都向应用程序日志发送事件记录信息。应用程序日志中记录的时间类型由应用程序的开发者决定，并提供相应的系统工具帮助用户使用应用程序日志。

（2）系统日志：记录由 Windows NT/2000 操作系统组件产生的事件，主要包括驱动程序、系统组件和应用软件的崩溃以及数据丢失错误等。系统日志中记录的时间类型由 Windows NT/2000 操作系统预先定义。

（3）安全日志：记录与安全相关事件，包括成功和不成功的登录或退出、系统资源使用事件等。与系统日志和应用程序日志不同，安全日志只有系统管理员才可以访问。

Windows NT/2000 的系统日志由事件记录组成。每个事件记录为 3 个功能区：记录头区、事件描述区和附加数据区。

2）日志文件默认位置

应用程序日志：%systemroot%\\system32\\config\\AppEvent.EVT。

系统日志：%systemroot%\\system32\\config\\SysEvent.EVT。

安全日志：%systemroot%\\system32\\config\\SecEvent.EVT。

FTP 日志：%systemroot%\\system32\\logfiles\\msftpsvc1\\。

WWW 日志：%systemroot%\\system32\\logfiles\\w3svc1\\。

Scheduler 服务日志：%systemroot%\\schedlgu.txt。

3）查看日志

查看日志有两种方法。

（1）选择“开始”→“设置”→“控制面板”→“管理工具”，在其中找到“事件查看器”。

（2）选择“开始”→“运行”，输入 eventvwr.msc，可以直接进入“事件查看器”。

4）设置日志保留方式

在 Windows 的“属性”对话框中，选中“定义这个策略设置”复选框，进行如下选择。

（1）“按需要改写事件”，选择将系统日志存档。

（2）“按天数改写事件”，设置合适的保存天数，但要确保系统日志足够大。

（3）“不要改写事件（手动清除日志）”，这种情况下，如果达到最大的日志大小，将丢弃新事件。

**4．UNIX 日志**

1）日志函数 syslog

UNIX 系统采用 syslog 函数记录日志。任何程序都可以通过 syslog 记录事件。syslog 可以记录系统事件，可以写到一个文件或设备中，或给用户发送一个信息。它能记录本地事件或通过网络记录另一台主机上的事件。

syslog 依据两个重要的文件——/sbin/syslogd（守护进程）和/etc/syslog.conf 配置文件。

2）日志子系统

在 Linux 系统中，有 3 个主要的日志子系统。

（1）连接时间日志：由多个程序执行，把记录写入到/var/log/wtmp 和/var/run/utmp，login 等程序更新 wtmp 和 utmp 文件，使系统管理员能够跟踪谁在何时登录到系统。

（2）进程统计日志：由系统内核执行。当一个进程终止时，为每个进程往进程统计文件（pacct 或 acct）中写一个记录。进程统计的目的是为系统中的基本服务提供命令使用统计。

（3）错误日志：由 syslogd 执行。各种系统守护进程、用户程序和内核通过 syslog 向文件/var/log/messages 报告值得注意的事件。另外有许多 UNIX 程序创建日志。像 HTTP 和 FTP 这样提供网络服务的服务器也保持详细的日志。

3）UNIX 常用日志文件

lastlog：记录用户最后一次成功登录的时间。

loginlog：不良的登录尝试记录。

messages：记录输出到系统主控台以及由 syslog 系统服务程序产生的消息。

utmp：记录当前登录的每个用户。

utmpx：扩展的 utmp。

wtmp：记录每一次用户登录和注销的历史信息。

wtmpx：扩展的 wtmp。

vold.log：记录使用外部介质出现的错误。

xferkig：记录 FTP 的存取情况。
sulog：记录 su 命令的使用情况。
acct：记录每个用户使用过的命令。
aculog：发出自动呼叫记录。
boot.log：记录开机启动信息。
secure：登录到系统存取资料的记录，如 FTP、SSH、Telnet 等。

4）日志文件的位置

在 UNIX 下，存放日志文件的常用目录如下。
（1）/usr/adm（早期版本的）。
（2）unix/var/adm（较新版本的）。
（3）unix/var/log（用于 Solaris、Linux 和 BSD 等）。

5）日志管理命令

UNIX 下日志查看和保留设置用命令方式进行，有兴趣者请查阅有关资料，这里不再介绍。

# 5.4 信息系统安全风险评估与审计

安全管理的第一步就是建立一个全局的安全目标，然后才能围绕这个总体目标指定系统的安全策略。但是，由于信息系统的重要性和激烈的攻防对抗，使得信息系统的脆弱性和面临的威胁不可避免，也使得人们不可能建立完全安全的、没有风险的信息系统。这里，风险就是脆弱性和威胁的总和。一个现实的目标则是，通过对于要保护的资产以及系统受到的潜在威胁的分析，把系统风险降低到可以接受的水平。这就是信息系统安全风险评估。

## 5.4.1 信息系统安全风险评估及其目的

### 1. 信息系统安全风险评估的概念

系统的安全强度可以通过风险大小衡量。科学地分析信息系统的风险，综合平衡风险和代价的过程就是信息系统安全风险评估。世界各国信息化的经验表明。
（1）不计代价，片面地追求系统安全是不切实际的。
（2）不考虑风险存在的信息系统是危险的，是要付出代价，甚至是灾难性代价的。
（3）所有的信息系统建设的生命周期都应当从安全风险评估开始。

### 2. 信息系统安全风险评估的目的

通过信息系统安全风险评估，组织可以达到如下目的。
（1）了解组织的信息系统的管理和安全现状。
（2）确定资产威胁源的分布，如入侵者、内部人员、自然灾害等；确定其实施的可能性；分析威胁发生后资产的价值损失、敏感性和严重性，确定相应级别；确定最敏感、最

重要资产在威胁发生后的损失。

（3）了解系统的脆弱性分布。

（4）明晰组织的安全需求，指导建立安全管理框架，合理规划安全建设计划。

**3. 信息系统安全风险评估时机**

信息系统安全风险评估是信息系统每个生命周期的起点和动因。具体地说，应当在下面的一些时机进行。

（1）要设计规划或升级到新的信息系统时。

（2）给目前的信息系统增加新的应用或新的扩充（包括进行互联）时。

（3）发生一次安全事件后。

（4）组织具有结构性变动时。

（5）按照规定或某些特殊要求对信息系统的安全进行评估时。

### 5.4.2 信息系统安全风险评估的准则与模式

**1. 信息系统安全风险评估准则**

在信息系统安全风险评估中，应当遵循如下一些原则。

（1）规范性原则。具有 3 层含义。

① 评估方案和实施，要根据有关标准进行。

② 选择的评估部门，需要被国家认可，并具有一定等级的资质。

③ 评估过程和文档要规范。

（2）整体性原则。评估要从业务的整体需求出发，不能局限于某些局部。

（3）最小影响原则。包含如下意义。

① 评估要有充分的计划性，不对系统运行产生显著影响。

② 所使用的评估工具要经过多次使用考验，具有很好的可控性。

（4）保密性原则。包含如下方面。

① 对评估数据严格保密。

② 不得泄露参评人员资料。

③ 不得使用评估数据对被评方造成利益损失。

**2. 信息系统安全风险评估参考标准**

进行信息系统安全风险评估时可以参照的标准如下。

① ASNZS 4306：1999（风险管理指南）——澳大利亚和新西兰关于风险管理的标准。

② NIST SP 800-30：美国国家标准和技术研究院（NIST）开发的信息系统风险管理指南。

③ NIST SP 800-26：NIST 开发的信息系统安全自我评估指南。

④ ISO 17799：英国标准协会（British Standard Institute，BSI）开发，后成为信息安全管理体系的国际标准。

⑤ BS 7799-2：BSI 开发的信息安全管理标准。

⑥ OCTAVE（Operationally Critical Threat，Asset，and Vulnerability Evaluation）：美国卡内基·梅隆大学软件工程学院开发的一种风险评估方法。

⑦ BS 15000（ITIL）：信息系统服务管理标准。

⑧ ISO 13335：信息技术安全管理指南。

⑨ G51：安全风险评估及审计指南。

⑩ ISO 15408/CC。

⑪ GB/T 18336：中国国家标准《信息技术、安全技术、信息技术安全性评估准则》。

⑫ GB 17859—1999：中国国家标准《计算机信息系统安全保护等级划分准则》。

**3. 信息系统安全风险评估模式**

安全风险评估模式是进行安全风险评估时应当遵循的操作过程和方式。每个组织应当根据自己的信息系统的环境选择适当的评估模式。下面是几种常用的风险评估模式。

1）基线评估（baseline risk assessment）

安全基线评估就是按照标准或惯例进行评估。例如，按照下列标准规范或者惯例。

（1）国际标准和国家标准，例如 BS 7799-1、GB/T 18336—2001 等。

（2）行业标准或推荐，例如德国联邦安全局 IT 基线保护手册等。

（3）其他具有类似商业目标和规模的组织的惯例。

采用基线安全风险评估，组织应当根据行业性质、业务环境等实际情况，用安全极限的规定对自己的信息系统的安全措施进行检查，找出差距，得到基本的安全需求。

安全基线规定适合于特定环境下的所有系统。采用基线安全风险评估，可以满足基本的安全需求，使系统达到一定强度的安全防护水平。这种评估模式需要的资源少，评估周期短，操作简单，是最经济有效的风险评估模式。但是，基线水平的确定较困难。

2）详细评估

详细评估就是对信息系统中的所有资源都进行仔细的评估。例如，可以划分成如下方面进行安全风险评估。

（1）网络安全风险评估，可以按照了解拓扑结构—获取公共访问机器名字和地址—进行端口扫描的顺序进行。

（2）平台安全风险评估，包括认证基准配置、操作系统、网络服务有无改变，认证管理员口令并测试口令的强度，跟踪审计子系统，评估数据库等。

（3）应用安全评估。

详细评估包括了资产的鉴定和评估、资产面临威胁的评估以及安全薄弱环节的分析，并在这些评估分析的基础上进行最后的风险评估分析，最后制定出合适的安全策略。它体现了风险管理的思想，能识别资产的风险并将风险降低到可以接受的水平。但是，这种模式需要相当多的财力、物力、时间、精力和专业能力的投入，最后获得的结果可能有一定的时间滞后。

3）组合评估

组合评估是上述两种模式的结合。它首先对所有信息系统进行一次较高级别的安全分析，并关注每一个实际分析对整个业务的价值以及它所面临的风险的程度。然后对鉴定为对业务非常重要或面临严重风险的部分，进行详细评估分析，对其他部分进行极限评估分析。这种方法注意了耗费与效率之间的平衡，还注意了高风险系统的安全防范。

### 5.4.3　信息系统安全风险评估过程

信息系统安全风险评估是确定信息系统安全需求的过程，它包括图 5.4 所示的几个阶段。

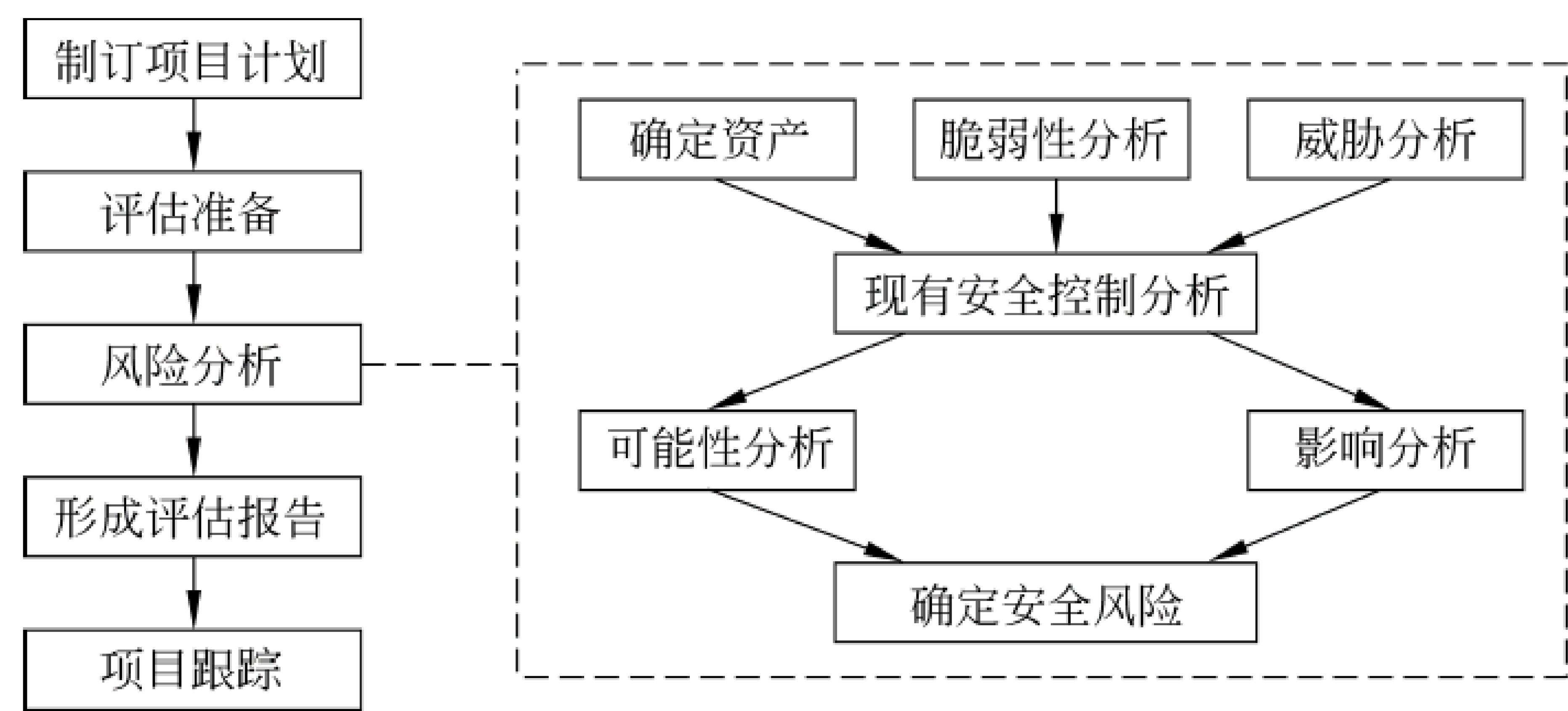



图 5.4　信息系统安全风险评估过程

下面对信息系统安全风险评估的各个阶段的工作作进一步说明。

**1. 制订项目计划**

评估工作从制订项目计划开始。项目计划应当包括如下一些内容。

（1）评估目标。进行安全风险评估的目的和期望。

（2）项目范围和边界。例如，通过定义系统的连接和接口。

（3）约束条件。包括时间（是否要在非繁忙办公时间，甚至非工作时间进行）、财务预算、技术因素等。这些约束可能影响项目进度和评估的可用资源。

（4）建立资产价值（重要性或敏感度）评估标准。

（5）风险接受标准。明确组织可以接受的风险的水平或等级。

（6）确定风险评估的模式。

（7）项目进度安排。用来控制进度，监督项目过程。

**2. 评估准备**

制订风险评估计划之后，便要为实施风险评估做准备工作。准备工作包括以下几个方面。

（1）成立一个专门的风险评估小组，成员包括：

① 具有风险评估经验者。

② 熟悉组织运作者。

③ 管理层、业务部门的成员。

④ IT 系统代表。

⑤ 用户代表。

⑥ 外部风险评估专家。

⑦ 组织的信息安全官员和安全管理人员。

⑧ 组织的高层管理人员。

同时要进行分工，明确责任。

（2）收集资料，围绕项目的范围，收集相关资料。收集方法包括：

① 问卷调查。

② 对各级人员进行访谈。

③ 小组讨论。

④ 查阅文档（政策法规、设计资料、操作指南、审计记录、安全策略、应急预案、以前的评估结果等）。

⑤ 现场勘察。观察各类人员的行为、环境状况和操作情形，寻找不良行为（如违反安全策略的现象）。

（3）材料准备。为风险评估过程设计拟定标准化的表格、模板和问卷等。

### 3. 风险分析

1）确定资产

通常用资产估计来确定需要保护的系统的价值。对于信息系统来说，可以用“信息资产”来描述信息化的成果。通常信息资产可以认为由组织的 5 种资产组成。

（1）物理资产。构成信息系统的一切具有物理形态的资产都称为物理资产，包括通信线路、通信设备、工作站、服务器、终端和存储设备等。

（2）信息资产。信息资产是相对于物理资产而言的资产，是具有信息属性的资产，包括软件以及各种信息资源（财务信息、人事信息、业务信息、计划信息、设计信息以及系统记录的其他信息等）。它们也常被看作是知识资产。

（3）时间资产。时间也是一种宝贵的资产。

（4）人力资源。人力资源是组织最灵活、最主动的资源。

（5）信誉（形象）资产。信誉是组织宝贵的无形资产。信誉受到损失，组织的形象和可信度将会不佳，愿意与之打交道者会减少。

资产的形式包括以下几种。

① 各种文档。包括数据库和数据文件、系统文件、用户手册、培训资料、支持程序和应急计划等。

② 纸质文件。包括合同、策略方针、企业文件和重要商业结果等。

③ 软件。应用软件、系统软件、开发工具和公用程序等。

④ 物理资产。

资产的确定应当从关键业务开始，最终覆盖所有关键资产。在实际操作时，可以根据关键业务流程确定资产清单。

得到完整的资产清单后，要进一步确定每项资产的价值。资产的价值用资产对于组织

的重要性或敏感度衡量。为了保证资产评估的一致性和准确性，组织应当建立一个资产评估标准，对资产进行等级划分。表 5.4 为一个资产敏感度等级划分标准范例。

**表 5.4　一个资产敏感度等级划分标准范例**

| 等级 | 名称 | 描　述 |
|---|---|---|
| 5 | 巨 | 造成灾难性损失，导致组织停顿，决策层免职 |
| 4 | 大 | 造成重大经济损失，造成产品和服务大幅度缩减，形象受损，士气低落 |
| 3 | 中 | 对组织造成引起重视的损失，市场有一定程度的反映，士气受到影响 |
| 2 | 小 | 对部分产品或服务出现影响，受到外部批评 |
| 1 | 微 | 出现影响，基本不产生负面效应 |

2）脆弱性分析

（1）脆弱性分类。

① 技术性脆弱性：系统软硬件中存在的漏洞或缺陷。

② 操作性脆弱性：系统在配置、操作、使用中的缺陷，包括操作人员的不良习惯、缺乏审计或备份等。

③ 管理性脆弱性：组织结构、人员意识、规章制度和策略计划等方面的不足。

（2）脆弱性分析手段。

① 技术性脆弱性分析手段。

- 采用工具进行网络扫描。
- 主机审计。采用脚本工具或人工方式，对网络设备、主机和数据库进行列目式排查。
- 渗透测试。人工模拟黑客攻击，进行排查。
- 系统分析。进行网络结构和边界分析。

② 非技术性脆弱性分析主要采用调查表、查看文件、访谈和现场勘察手段进行。

3）威胁分析

威胁是对系统或资产的保密性、完整性以及可用性构成潜在损害的事件。威胁分析的目的是在明晰关键资产安全需求的基础上，确定面临的威胁，并界定发生威胁的可能性以及对系统或资产的破坏性潜力。

（1）威胁源分类。

为了描述方便，对威胁源要进行分类。表 5.5 为一个威胁源分类范例。

**表 5.5　一个威胁源分类范例**

| 编号 | 名　称 | 描　述 |
|---|---|---|
| 1 | 不可抗 | 不可抗拒的自然灾害（地震、飓风等）、环境（电力中断、污染等）、政治、战争等 |
| 2 | 组织薄弱 | 因组织、体制或制度缺陷造成的安全威胁 |
| 3 | 人为失误 | 因人的素质、技能等形成的安全威胁 |
| 4 | 技术缺陷 | 因技术缺陷形成的安全威胁 |
| 5 | 恶意行为 | 人为的侵害行为 |

（2）威胁确定途径。

威胁确定可以通过下列途径进行。

① 查看安全策略文档。

② 业务流程分析。

③ 网络拓扑分析。

④ 人员访谈。

⑤ 入侵检测系统收集信息分析。

⑥ 人工分析。

⋮

4）现有安全控制分析

对于现有（在规划中的或已经实现的）安全控制措施进行分析的目的，是通过分析这些控制措施减少或消除一个威胁源利用系统脆弱性的可能性。

（1）安全控制措施的类型。

安全控制措施按照性质分为以下几种。

① 管理性（administrative）。包括安全策略、程序管理、风险管理、安全保障和系统生命周期管理等。

② 操作性（operational）。包括人员职责、应急响应、事件处理、意识教育、系统支持和操作、物理和环境安全等。

③ 技术性（technical）。加密、认证、访问控制和审计等。

安全控制措施按照功能分为以下几种。

① 预防性（preventive）。阻止对安全策略的犯罪，包括访问控制、加密和认证等。

② 检测性（detective）。检测并及时发现对安全策略的违犯或企图，并发出警告，具有一定威慑性（deterrent），如入侵检测、审计跟踪、校验和、蜜罐技术等。

（2）安全控制措施分析方法。

设计一个安全要求核对表，系统化地进行有效分析，验证安全是否与既定法规和政策一致。

（3）结果。

输出信息系统已经实现或计划实现的安全控制措施清单。

5）可能性及影响分析

可能性（likelihood）和影响（impact）是威胁的两个属性。也是评估风险的两个关键因素。可能性指威胁发生的概率，影响指用于确定风险发生威胁对系统资产破坏或影响的程度。

（1）可能性分析。

可能性分析是对威胁发生的概率的估计，要结合威胁源的动机和能力、脆弱性的性质、安全控制措施的存在与有效性进行综合评估。通常采用经验分析或定性分析的方法确

定。为便于分析，应当制定一个威胁的可能性等级标准。表 5.6 为一个威胁的可能性等级范例。

**表 5.6　一个威胁的可能性等级范例**

| 等级 | 名称 | 等级权重 | 描　　述 |
|---|---|---|---|
| A | 频繁 | 1.0 | 大多数情况下会发生 |
| B | 经常 | 0.7 | 多数情况下很可能发生 |
| C | 有时 | 0.5 | 有时会发生 |
| D | 很少 | 0.3 | 有时可能发生 |
| E | 个别 | 0.1 | 特殊情况下发生 |

（2）影响分析。

影响分析已经在确定资产阶段用资产的敏感性进行了描述。需要说明的是，影响分析可以用定性和定量两种方法进行。

定性影响分析只用级别描述威胁的影响，这样可以对风险进行排序，并能够立即对那些需要改善的环节进行标识。定量分析可以计算影响的大小，以便用成本效益分析进行成本控制。

6）确定安全风险

确定安全风险的目的是评估信息系统的安全风险级别。

（1）风险级别和措施。

信息系统风险级别最多划分为 4 级，并可以用颜色表示，如表 5.7 所示。

**表 5.7　信息系统风险级别和行动措施**

| 级别符号/颜色 | 名　称 | 建议的行动措施 |
|---|---|---|
| E/红色 | 极度风险 | 立即采取措施：避免、转移、降低 |
| H/橙色 | 高风险 | 需要尽快部署行动：避免、转移、降低 |
| M/黄色 | 中风险 | 必须在一个合理的时间段内制订一个计划实施行动：避免、接受、转移、降低 |
| L/绿色 | 低风险 | 按常规处理：避免、接受、转移、降低 |

（2）风险级别矩阵。

将风险的可能性（概率）与威胁的级别相乘，可以得到最终使命风险，从而可以得到风险矩阵。表 5.8 为一个计算范例。

（3）确定风险尺度。

例如，对风险级别作如下定义：E 为 4.5～5.0，H 为 3.5～4.5，M 为 1.0～3.5，L 为 0.1～1.0。

按照风险级别的上述定义，将风险级别符号标进风险矩阵，得到的结果如表 5.9 所示。

表 5.8　一个风险矩阵计算范例

| 可 能 性 | 影　　响 | | | | |
|---|---|---|---|---|---|
| | 微 | 小 | 中 | 大 | 巨 |
| | 1 | 2 | 3 | 4 | 5 |
| A（1.0） | 1.0 | 2.0 | 3.0 | 4.0 | 5.0 |
| B（0.7） | 0.7 | 1.4 | 2.1 | 2.8 | 3.5 |
| C（0.5） | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 |
| D（0.3） | 0.3 | 0.6 | 0.9 | 1.2 | 1.5 |
| E（0.1） | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 |

表 5.9　一个风险矩阵结果

| 可 能 性 | 影　　响 | | | | |
|---|---|---|---|---|---|
| | 微 | 小 | 中 | 大 | 巨 |
| | 1 | 2 | 3 | 4 | 5 |
| A（1.0） | M（1.0） | M（2.0） | M（3.0） | H（4.0） | E（5.0） |
| B（0.7） | L（0.7） | M（1.4） | M（2.1） | M（2.8） | H（3.5） |
| C（0.5） | L（0.5） | M（1.0） | M（1.5） | M（2.0） | M（2.5） |
| D（0.3） | L（0.3） | L（0.6） | M（0.9） | M（1.2） | M（1.5） |
| E（0.1） | L（0.1） | L（0.2） | L（0.3） | L（0.4） | L（0.5） |

7）形成评估报告

风险评估报告内容一般包括以下几部分。

（1）概述：评估目的、方法和过程等。

（2）评估结果：包括资产、威胁、脆弱性、现有安全控制措施等级和风险评估等。

（3）安全控制建议和备选解决方案。

前两项的内容已经介绍过了，在对安全控制建议和备选解决方案提出建议时应当考虑的内容如下。

① 建议的选项在兼容性等方面的有效性。

② 与法律法规的符合性。

③ 组织及策略方面的可接受性。

④ 对运行的影响。

⑤ 安全性和可靠性。

### 5.4.4　信息系统渗透测试

一般说来，渗透测试（penetration test）是一种通过模拟恶意攻击者的技术与方法，挫败目标系统的安全控制措施，取得访问控制权，并发现具备业务影响后果安全隐患的一种安全测试与评估方式。这种方法源于军事演习，20 世纪 90 年代时，由美国军方与国家安全局引入到对信息网络与信息安全基础设施的实际攻防测试过程中。

实施渗透测试一般需要对目标系统进行主动探测分析，以发现潜在的系统漏洞，包括不

恰当的系统配置、已知或未知的软硬件漏洞以及在安全计划与响应过程中的操作性弱点等。

### 1. 渗透测试方法

1）黑箱测试

黑箱测试（black-box testing）又称为 zero-knowledge testing。采用这种方式时，渗透测试团队将从一个远程网络位置来评估目标网络基础设施，并没有任何目标网络内部拓扑等相关信息，完全处于对系统一无所知的状态。他们完全模拟真实网络环境中的外部攻击者，采用流行的攻击技术与工具，有组织有步骤地对目标组织进行逐步的渗透与入侵，揭示目标网络中一些已知或未知的安全漏洞，并评估这些漏洞能否被用来获取控制权或造成业务资产的损失。所以这种测试方法又称为外部测试（external testing）。

黑盒测试还可以对目标组织内部安全团队的检测与响应能力做出评估。在测试结束之后，黑盒测试会对发现的目标系统安全漏洞、所识别的安全风险及其业务影响评估等信息进行总结和报告。

2）白盒测试

白盒测试（white-box testing）也称为内部测试（internal testing）。进行白盒测试需要事先了解关于目标环境的所有内部与底层状况，因此可以让渗透测试者以最小的代价发现和验证系统中最严重的安全漏洞。如果实施到位，白盒测试能够比黑盒测试消除更多的目标基础设施环境中的安全漏洞与弱点，从而给客户组织带来更大的价值。这类测试通常用于模拟企业内部雇员的越权操作。

3）灰盒测试

灰盒测试（grey-box testing）是对黑盒测试和白盒测试两种基本类型的组合。采用灰盒测试可以提供对目标系统更加深入和全面的安全审查，能够同时发挥两种基本类型渗透测试方法的各自优势。这种测试也称为隐秘测试，是被测单位仅有极少数人知晓测试存在的方法，因此能够有效地检验单位中的信息安全事件监控、响应、恢复做得是否到位。

### 2. 渗透目标

1）主机操作系统渗透

对 Windows、Linux 等操作系统本身进行渗透测试。

2）数据库系统渗透

对 MS-SQL、Oracle、MySQL 等数据库应用系统进行渗透测试。

3）应用系统渗透

对渗透目标提供的各种应用，如 ASP、JSP、PHP 等组成的 WWW 应用进行渗透测试。

4）网络设备渗透

对各种防火墙、入侵检测系统、网络设备进行渗透测试。

5）不同网段/VLAN 之间的渗透

这种渗透方式是从某内/外部网段，尝试对另一网段/VLAN 进行渗透。这类测试通常可能用到的技术包括对网络设备的远程攻击、对防火墙的远程攻击或规则探测以及规避尝试。

### 3. 渗透测试手段

1）内网测试

内网测试指的是渗透测试人员由内部网络发起测试，这类测试能够模拟企业内部违规操作者的行为。最主要的“优势”是绕过了防火墙的保护。内部主要可能采用的渗透方式包括远程缓冲区溢出、口令猜测以及 B/S 或 C/S 应用程序测试（如果涉及 C/S 程序测试，需要提前准备相关客户端软件供测试使用）。

2）外网测试

外网测试指的是渗透测试人员完全处于外部网络（例如拨号、ADSL 或外部光纤），模拟对内部状态一无所知的外部攻击者的行为。包括对网络设备的远程攻击，口令管理安全性测试，防火墙规则试探和规避，Web 及其他开放应用服务的安全性测试。

3）端口扫描

通过对目的地址的 TCP/UDP 端口扫描，确定其所开放的服务的数量和类型，这是所有渗透测试的基础。通过端口扫描，可以基本确定一个系统的基本信息，结合安全工程师的经验可以确定其可能存在以及被利用的安全弱点，为进行深层次的渗透提供依据。

4）远程溢出

这是当前出现的频率最高、威胁最严重，同时又是最容易实现的一种渗透方法，一个具有一般网络知识的入侵者就可以在很短的时间内利用现成的工具实现远程溢出攻击。

对于防火墙内的系统同样存在这样的风险，只要对跨接防火墙内外的一台主机攻击成功，那么通过这台主机对防火墙内的主机进行攻击就易如反掌。

5）本地溢出

本地溢出是指在拥有了一个普通用户的账号之后，通过一段特殊的指令代码获得管理员权限的方法。使用本地溢出的前提是首先要获得一个普通用户密码。也就是说，导致本地溢出的一个关键条件是设置不当的密码策略。

多年的实践证明，在经过前期的口令猜测阶段获取的普通账号登录系统之后，对系统实施本地溢出攻击，就能获取不进行主动安全防御的系统的控制管理权限。

6）口令猜测

口令猜测也是一种出现概率很高的风险，几乎不需要任何攻击工具，利用一个简单的暴力攻击程序和一个比较完善的字典就可以猜测口令。

对一个系统账号的猜测通常包括两个方面：首先是对用户名的猜测，其次是对密码的猜测。

7）脚本及应用测试

Web 脚本及应用测试专门针对 Web 及数据库服务器进行。根据最新的技术统计，脚本安全弱点为当前 Web 系统，尤其是存在动态内容的 Web 系统比较严重的安全弱点之一。利用脚本相关弱点轻则可以获取系统其他目录的访问权限，重则将有可能取得系统的控制权限。因此对于含有动态页面的 Web、数据库等系统，Web 脚本及应用测试将是必不可少的一个环节。在 Web 脚本及应用测试中，可能需要检查的部分包括如下。

（1）检查应用系统架构，防止用户绕过系统直接修改数据库。

（2）检查身份认证模块，防止非法用户绕过身份认证。

（3）检查数据库接口模块，防止用户获取系统权限。

（4）检查文件接口模块，防止用户获取系统文件。

（5）检查其他安全威胁。

8）无线网络测试

通过对无线网络的测试，可以判断企业局域网的安全性，已经成为越来越重要的渗透测试环节。

除了上述测试手段外，还有一些可能会在渗透测试过程中使用的技术，包括社交工程学、拒绝服务攻击以及中间人攻击。

**4. 渗透测试流程**

渗透测试一般包括如下 7 个阶段。

1）前期交互阶段

在前期交互（pre-engagement interaction）阶段，渗透测试团队要与客户组织进行交互讨论，收集客户需求，准备测试计划，定义测试范围、边界和限制条件，定义业务目标、项目管理与规划，讨论服务合同细节等。

2）情报搜集阶段

情报搜集（information gathering）是在目标范围确定之后，进行尝试获取更多关于目标组织网络拓扑、系统配置与安全防御措施信息的活动。

对目标系统的情报探查能力是渗透测试者的一项非常重要的技能，情报搜集是否充分在很大程度上决定了渗透测试的成败。所以在这个阶段，渗透测试团队要尽力利用各种信息来源与搜集技术方法，包括公开来源信息查询、Google Hacking、社会工程学、网络踩点、扫描探测、被动监听和服务查点等。

3）威胁建模阶段

在搜集到充分的情报信息后，渗透测试团队将集思广益，通过缜密的情报分析，进入威胁建模（threat modeling）、制定攻击规划阶段。

4）漏洞分析阶段

漏洞分析（vulnerability analysis）阶段的工作是在确定出最可行的攻击通道之后，在前

几个阶段获取的情报信息，特别是在安全漏洞扫描结果、服务查点信息等基础上，通过搜索可获取的渗透代码资源，找出可以实施渗透攻击的攻击点，并在实验环境中进行攻击模拟。高水平的渗透测试团队还会针对攻击通道上的一些关键系统与服务进一步进行安全漏洞探测与挖掘，找出可被利用的未知安全漏洞，并开发出渗透代码，打开攻击通道上的关键路径。

在这个环节中，需要渗透测试团队根据目标组织的业务经营模式、保护资产形式与安全防御计划的不同特点，自主设计出攻击目标，识别关键基础设施，并寻找客户组织最具价值和尝试安全保护的信息和资产，最终达成能够对客户组织造成最重要业务影响的攻击途径。

5）渗透攻击阶段

在渗透攻击（exploitation）将实施真正的入侵，获得访问控制权。

6）后渗透攻击阶段

在后渗透攻击（post exploitation）阶段，渗透测试团队要根据测试目的、收集的信息和测试结果，对系统的安全状况进行科学的评价，并评估自己的渗透攻击，查找不足。若有不足或遗漏，应返回到威胁建模阶段进行弥补。

7）渗透测试报告撰写阶段

渗透测试报告（reporting）是提交给用户的最终成果，内容包括所有阶段中渗透测试团队所获取的关键情报信息、探测和发掘出的系统安全漏洞、成功渗透攻击的过程，以及造成业务影响后果的攻击途径，同时还要站在防御者的角度，帮助客户分析安全防御体系中的薄弱环节以及存在的问题，给出系统加固建议。

通常也把前 4 个阶段称为预攻击阶段，把第 5 个阶段称为预攻击阶段，把最后两个阶段称为后攻击阶段。

**5. 渗透测试的实施**

渗透测试是按照黑客攻击的思路进行的攻击性测试，每个阶段、对于不同的部位的攻击，都可以采用已有的攻击工具进行。图 5.5 给出实施渗透测试所进行的有关攻击的目标及其使用的工具。

### 5.4.5　信息系统安全审计

**1. 信息系统安全审计及其功能**

审计（audit）是按照一定的标准对于一个过程所进行的监督和评价机制。审计的执行是以确定有效性和可靠性的信息为依据，以检查、验证目标的准确性和完整性为目的。信息系统审计是一个通过收集和评价审计证据，对信息系统是否能够保护资产的安全、维护数据的完整、使被审计单位的目标得以有效地实现、使组织的资源得到高效地使用等方面做出判断的过程，其目标是协助组织的信息技术管理人员有效地履行其责任，以达成组织的信息技术管理目标。组织的信息技术管理目标是保证组织的信息技术战略，充分反映组

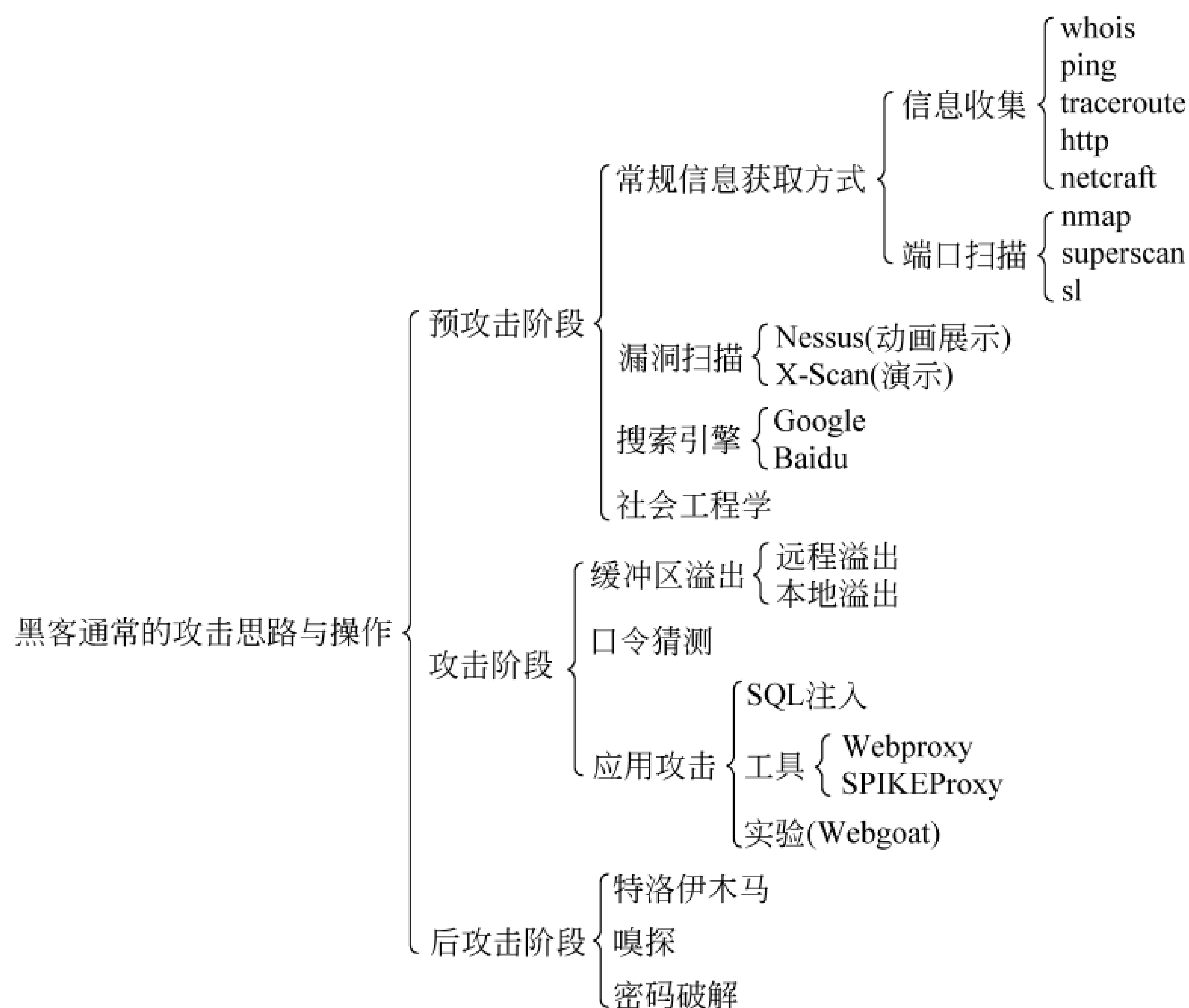



图 5.5　实施渗透测试所进行的有关攻击目标及其使用的工具

织的业务战略目标，提高组织所依赖的信息系统的可靠性、稳定性、安全性及数据处理的完整性和准确性，提高信息系统运行的效果与效率，保证信息系统的运行符合法律、法规及监管的相关要求。

国际通用的 CC 准则（即 ISO/IEC 15408-2:1999《信息技术安全性评估准则》）中对信息系统安全审计（Information System Security Audit，ISSA）给出了明确定义：信息系统安全审计主要指对与安全有关的活动的相关信息进行识别、记录、存储和分析；审计记录的结果用于检查网络上发生了哪些与安全有关的活动，谁（哪个用户）对这个活动负责；主要功能包括安全审计自动响应、安全审计数据生成、安全审计分析、安全审计浏览、安全审计事件选择和安全审计事件存储等。

简单地说，安全审计应当具有下面的功能。

（1）记录关键事件。关于关键事件的界定由安全官员决定。

（2）对潜在的攻击者进行威慑或警告。

（3）为系统安全管理员提供有价值的系统使用日志，帮助系统管理员及时发现入侵行为和系统漏洞，使安全管理人员可以知道如何对系统安全进行加强和改进。

（4）为安全官员提供一组可供分析的管理数据，用于发现何处有违反安全方案的事件，并可以根据实际情形调整安全政策。

### 2. 信息系统安全审计的基本内容

1）组织控制审计

组织控制审计包括如下内容。

（1）了解被审计单位的组织结构、人员分工、业务授权和职责分离等情况。

（2）审查组织控制措施是否健全，是否制定了完善的工作制度和岗位职责，是否落实了岗位责任制和风险防范责任，是否建立了内部监督机制和考核机制等。

2）安全控制审计

安全控制审计包括如下内容。

（1）环境控制审计，包括是否为信息系统的硬件设备提供适合的工作环境，保证设备正常运转。

（2）技术安全控制审计，包括是否通过加密技术限制未经授权的人员接触机密数据和文件，是否有系统硬软件和数据文件的灾难补救计划，是否定期或在重要操作前对数据进行备份，以减少意外导致损失的可能性。

3）部位安全审计

部位安全审计包括以下几种审计。

（1）系统设备的安全审计。

（2）操作系统安全的审计。

（3）网络及其应用的安全审计。

（4）数据库系统安全审计。

（5）应用系统的安全审计。

（6）环境安全审计。

4）信息系统威胁审计

信息系统威胁审计包括以下几种审计。

（1）对来自外部攻击的审计。

（2）对来自内部攻击的审计。

5）对电子数据的安全审计

**3. 信息系统安全审计的基本步骤**

（1）编制组织使用的信息系统清单并对其进行分类。

（2）决定哪些系统影响关键功能和资产。

（3）评估哪些风险影响这些系统及对商业运作的冲击。

（4）在上述评估的基础上对系统分级，决定审计优先值、资源、进度和频率。审计者可以制订年度审计计划，开列出一年之中要进行的审计项目。

**4. 信息系统安全审计工具**

审计可以采用如下一些工具。

（1）检验表：是为审计工作标准化提供的一种简捷的工具。主要分为 3 类。

① 审计检验表。可以分为被审部门校验（检查）表（分为审计项目、审计方法和审计结论列表）和审计条款校验（检查）表（分为责任部门、审计内容、检查方式和审计结论

列表）等。

② 设置检验表。

③ 漏洞检验表。

（2）扫描工具：进行 IP、端口号和漏洞扫描。

（3）完整性检验工具：进行完整性保护检验，如 Triwire 等。

（4）渗透测试工具。

# 5.5　信息系统安全测评准则

任何信息系统的安全需求都不是无限的，因为无限的需求需要无限的投入。因此对于信息系统的安全确立一个评价准则非常必要。

## 5.5.1　国际信息安全测评准则

世界上最早的计算机系统安全标准是美国国防部于 1979 年 6 月 25 日发布的军标 DoD 5200.28-M。在此基础上，美国国防部于 1983 年发布可信计算机系统评价准则 TCSEC（1985 年发布正式版），又称为橘皮书。以后，许多国家和国际组织也相继提出了新的安全评价准则。图 5.6 所示为国际主要信息技术安全测评标准的发展及其联系。

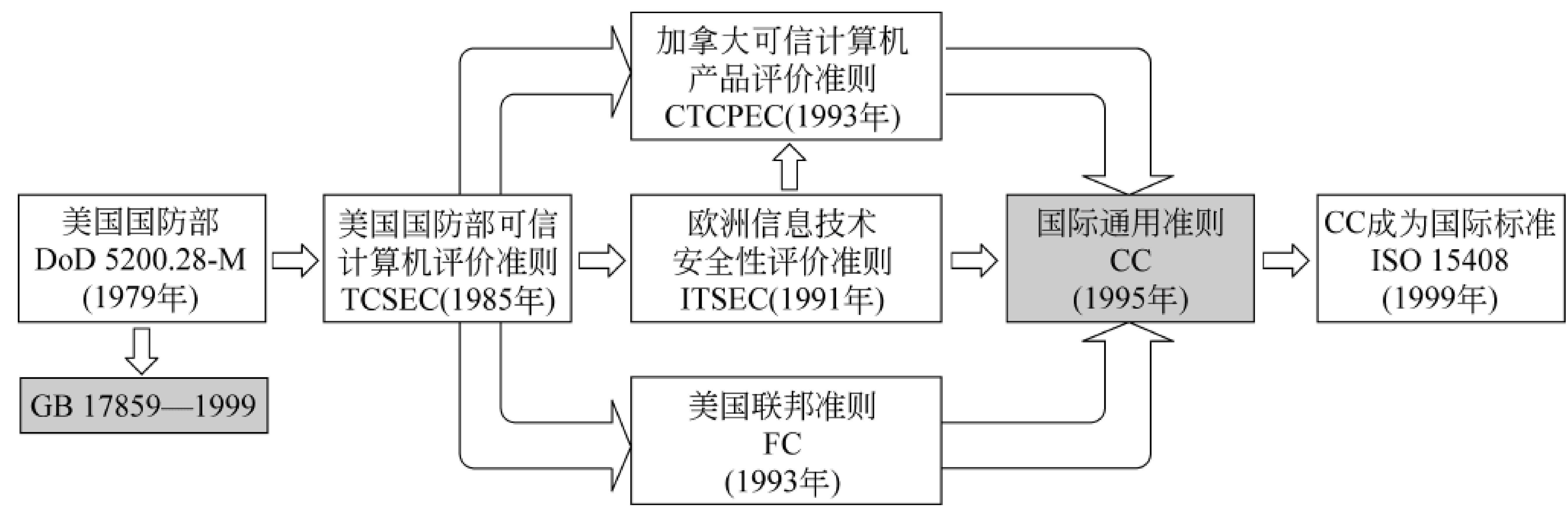



图 5.6　国际主要信息技术安全测评标准的发展及其联系

在信息安全等级标准中，一个非常重要的概念是可信计算基（Trusted Computer Base，TCB）。TCB 是计算机系统赖以实施安全性的一切设施，包括硬件、固件、软件和负责安全策略的组合。它们根据安全策略来处理主体（系统管理员、安全管理员、用户和进程）对客体（进程、文件、记录和设备等）的访问，通常包括下列部分。

（1）操作系统的安全内核。

（2）具有特权的程序和命令。

（3）处理敏感信息的程序，如系统管理命令等。

（4）与 TCB 实施安全策略有关的文件。

（5）其他有关的固件、硬件和设备。

（6）负责系统管理的人员。

（7）保障固件和硬件正确的程序和诊断软件。

（8）具有抗篡改的性能和易于分析与测试的结构。

### 1．DoD 5200.28-M

DoD 5200.28-M 为计算机系统定义了 4 种不同的运行模式。

（1）受控的安全模式：系统用户对系统的机密资料的访问控制没有在操作系统中实现，安全的实现可以通过控制用户对计算机的操作权限等管理措施实现。

（2）自主安全模式：计算机系统和外围设备可以在指定用户或用户群的控制下工作，该类用户了解并可自主地设置机密资料的类型与安全级别。

（3）多级安全模式：系统允许不同级别和类型的机密资料并存和并发处理，并且有选择地许可不同的用户对存储数据进行访问。用户与数据的隔离控制由操作系统和相关系统软件实现。

（4）强安全模式：所有系统部件依照最高级别类型得到保护，所有系统用户必须有一个安全策略；系统的控制操作对用户透明，由系统实现对机密资料的并发控制。

### 2．TCSEC

TCSEC 是计算机系统安全评价的第一个正式标准，于 1970 年由美国国防科学技术委员会提出，于 1985 年 12 月由美国国防部公布。TCSEC 把计算机系统的安全分为如下 4 等 7 级。

1）D 等（含 1 级）

D1 级系统：最低级。只为文件和用户提供安全保护。

2）C 等（含 2 级）

C1 级系统：可信计算基通过用户和数据分开来达到安全目的，使所有的用户都以同样的灵敏度处理数据（可认为所有文档有相同的机密性）。

C2 级系统：在 C1 级的基础上，通过登录、安全事件和资源隔离增强可调的审慎控制。在连接到网上时，用户分别对自己的行为负责。

3）B 等（含 3 级）

B 级具有强制性保护功能。强制性意味着在没有与安全等级相连的情况下，系统就不会让用户存取对象。

B1 级系统要求如下。

（1）对每个对象都进行灵敏度标记，导入非标记对象前要先标记它们。

（2）用灵敏度标记作为强制访问控制的基础。

（3）灵敏度标记必须准确地表示其所联系的对象的安全级别。

（4）系统必须使用用户口令或身份认证来决定用户的安全访问级别。

（5）系统必须通过审计来记录未授权访问的企图。

B2 级系统要求如下。

（1）必须符合 B1 级系统的所有要求。

（2）系统管理员必须使用一个明确的、文档化的安全策略模式作为系统可信任运算基础体制；可信任运算基础体制能够支持独立的操作者和管理员。

（3）只有用户能够在可信任通信路径中进行初始化通信。

（4）所有与用户相关的网络连接的改变必须通知所有的用户。

B3 级系统具有很强的监视委托管理访问能力和抗干扰能力。要求如下。

（1）必须符合 B2 级系统的所有安全需求。

（2）必须设有安全管理员。

（3）除控制个别对象的访问外，必须产生一个可读的安全列表；每个被命名的对象提供对该对象没有访问的用户列表说明。

（4）系统验证每一个用户身份，并会发送一个取消访问的审计跟踪消息。

（5）设计者必须正确区分可信任路径和其他路径。

（6）可信任的通信基础体制为每一个被命名的对象建立安全审计跟踪。

（7）可信任的运算基础体制支持独立的安全管理。

4）A 等（只含 1 级）

A1 级：最高安全级别。A1 级与 B3 级相似，对系统的结构和策略不作特别要求，而系统的设计者必须按照一个正式的设计规范进行系统分析；分析后必须用核对技术确保系统符合设计规范。A1 级系统必须满足如下要求。

（1）系统管理员必须接收到开发者提供的安全策略正式模型。

（2）所有的安装操作都必须由系统管理员进行。

（3）系统管理员进行的每一步安装操作必须有正式的文档。

TCSEC 的初衷主要是针对集中式计算的分时多用户操作系统。后来又针对网络（分布式）和数据库管理系统（C/S 结构）补充了一些附加说明和解释，典型的有可信计算机网络系统说明（NCSC-TG-005）和可信数据库管理系统解释等。

### 3. 欧共体信息技术安全评价准则 ITSEC

ITSEC 是欧共体于 1991 年发布的，它是欧洲多国安全评价方法的综合产物，应用领域为军队、政府和商业。该标准将安全的概念分为功能和评估两部分。

1）功能准则

分为 10 级：F1～F10。

（1）F1～F5 对应 TCSEC 的 C1 等～A 等。

（2）F6～F10 对应数据和程序的完整性、系统的可用性、数据通信的完整性和保密性。

2）评估准则

分为 6 级，分别是测试、配置控制和可控的分配、详细设计和编码、详细的脆弱性分析、设计与源代码明显对应以及设计与源代码在形式上的一致。

### 4. 加拿大可信计算机产品安全评价准则 CTCPEC

CTCPEC 是加拿大于 1993 年发布的。它综合了 TCSEC 和 ITSEC 两个准则的优点，专门针对政府需求设计。它将安全分为功能性需求和保证性需求两部分。功能性需求分为 4 大类。

（1）机密性。

（2）可用性。

（3）完整性。

（4）可控性。

每一类安全需求又分为一些小类（分级条数 0～5），以表示安全性上的差别。

### 5. 美国信息技术安全评价联邦准则 FC

FC 也吸收了 TCSEC 和 ITSEC 两个准则的优点，于 1993 年发布的。它引入了“保护轮廓”（PP）的概念。每个轮廓都包括功能、开发保证和评价 3 部分，在美国政府、民间和商业上应用很广。

### 6. 国际通用准则 CC

1993 年 6 月，欧美的 6 个国家将各自独立的准则集合成一系列单一的、能被广泛接受的 IT 安全准则——通用准则（CC），将 CC 提交给 ISO，于 1996 年颁布了 1.0 版。1999 年 12 月 ISO 正式将 CC 2.0（1998 年颁布）作为国际标准——ISO 15408 发布。

CC 的主要思想和框架都取自 ITSEC 和 FC，并突出了“保护轮廓”的概念，将评估过程分为安全保证和安全功能两部分。安全保证分为 7 个评估保证级别。

EAL1：功能测试。

EAL2：结构测试。

EAL3：系统测试和检查。

EAL4：系统设计、测试和复查。

EAL5：半形式化设计和测试。

EAL6：半形式化验证的设计和测试。

EAL7：集成化验证的设计和测试。

表 5.10 为 CC、TCSEC 和 ITSEC 标准的对应关系。

**表 5.10　CC、TCSEC 和 ITSEC 标准的对应关系**

| 标　准 | 等　　级 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| CC | — | EAL1 | EAL2 | EAL3 | EAL4 | EAL5 | EAL6 | EAL7 | |
| TCSEC | D | | C1 | C2 | B1 | B2 | B3 | A | 超 A |
| ITSEC | E0 | | E1 | E2 | E3 | E4 | E5 | E6 | |

CC 目前已经发布了如下版本。

（1）1996 年 6 月发布 CC 第 1 版。

（2）1998 年 5 月发布 CC 第 2 版。

（3）1999 年 10 月发布 CC 第 2.1 版，并成为 ISO 标准。

### 5.5.2 中国信息系统安全保护等级划分准则

中国已经发布实施《计算机信息系统安全保护等级划分准则》GB 17859—1999。这是一部强制性国家标准，也是一种技术法规。它是在参考了 DoD 5200.28-STD 和 NCSC-TC-005 的基础上，从自主访问控制、强制访问控制、标记、身份鉴别、客体重用、审计、数据完整性、隐蔽信道分析、可信路径和可恢复 10 个方面将计算机信息系统安全保护等级划分为 5 个级别的安全保护能力。

第一级：用户自主保护级，相当于 TCSEC 中定义的 C1 级。

第二级：系统审计保护级，相当于 TCSEC 中定义的 C2 级。

第三级：安全标记保护级，相当于 TCSEC 中定义的 B1 级。

第四级：结构化保护级，相当于 TCSEC 中定义的 B2 级。

第五级：访问验证保护级，相当于 TCSEC 中定义的 B3 级。

计算机信息系统的安全保护能力随着安全保护等级的提高而增强。

在信息安全等级标准中，各等级之间的差异在于 TCB 的构造不同以及其所具有的安全保护能力的不同。表 5.11 为这 5 个级别之间的简单比较。

**表 5.11　5 个级别之间的比较**

| | 第一级<br>用户自主保护级 | 第二级<br>系统审计保护级 | 第三级<br>安全标记保护级 | 第四级<br>结构化保护级 | 第五级<br>访问验证保护级 |
|---|---|---|---|---|---|
| 自主访问控制 | • | • | • | • | • |
| 身份鉴别 | • | • | • | • | • |
| 数据完整性 | • | • | • | • | • |
| 客体重用 | | • | • | • | • |
| 审计 | | • | • | • | • |
| 强制访问控制 | | | • | • | • |
| 标记 | | | • | • | • |
| 隐藏信道分析 | | | | • | • |
| 可信路径 | | | | • | • |
| 可信恢复 | | | | | • |

下面介绍各等级的基本内容。

#### 1. 第一级：用户自主保护级

本级可信计算基（TCB）通过隔离用户与数据，使用户具备自主安全保护的能力。它具有多种形式的控制能力，对用户实施访问控制，即为用户提供可行的手段，保护用户和用户组信息，避免其他用户对数据的非法读写与破坏。具体保护能力如下。

（1）自主访问控制：可信计算基定义系统中的用户和命名用户对命名客体的访问，并允许命名用户以自己的身份和（或）用户组的身份指定并控制对客体的访问；阻止非授权

用户读取敏感信息。

（2）身份鉴别：从用户的角度看，可信计算基的责任就是进行身份鉴别。在系统初始化时，首先要求用户标识自己的身份，并使用保护机制（例如口令）来鉴别用户的身份，阻止非授权用户访问用户身份鉴别数据。

（3）数据完整性：可信计算基通过自主完整性策略，阻止非授权用户修改或破坏敏感信息。

### 2. 第二级：系统审计保护级

这一级除具备第一级所有的安全功能外，要求创建和维护访问的审计跟踪记录，使所有用户对自己的合法性行为负责。具体保护能力如下。

（1）自主访问控制：可信计算基定义实施的访问控制的粒度是单个用户。没有存取权的用户只允许由授权用户指定对客体的访问权。

（2）身份鉴别比用户自主保护级增加两点。

① 通过为用户提供唯一标识，可信计算基使用户对自己的行为负责。

② 具备将身份标识与该用户所有可审计行为相关联的能力。

（3）客体重用：在可信计算基的空闲存储客体空间中，对客体初始指定、分配或再分配一个主体之前，撤销该客体所含信息的所有授权。当主体获得对一个已被释放的客体的访问权时，当前主体不能获得原主体活动所产生的任何信息。

（4）审计：在可信计算基能创建和维护受保护客体的访问审计跟踪记录，并能阻止非授权的用户对它访问或破坏。

可信计算基能记录下述事件：使用的身份鉴别机制；将客体引入用户地址空间（例如打开文件、程序初始化）；删除客体；由操作员、系统管理员或（和）系统安全管理员实施的动作，以及其他与系统安全有关的事件。对于每一事件，其审计记录包括事件的日期和时间、用户、事件类型、事件是否成功。对于身份鉴别事件，审计记录包含请求的来源（例如终端标识符）；对于客体引入用户地址空间的事件及客体删除事件，审计记录包含客体名。对不能由可信计算基独立分辨的审计事件，审计机制提供审计记录接口，可由授权主体调用。这些审计记录区别于计算机信息系统可信计算基独立分辨的审计记录。

（5）数据完整性：可信计算基通过自主完整性策略，阻止非授权用户修改或破坏敏感信息。

### 3. 第三级：安全标记保护级

本级的可信计算基具有系统审计保护级的所有功能。此外，还需以访问对象的安全级别限制访问者的访问权限，实现对访问对象的强制访问。为此需要提供有关安全策略模型、数据标记以及主体对客体强制访问控制的非形式化描述，具有准确地标记输出信息的能力，消除测试发现的任何错误。具体保护能力如下。

（1）自主访问控制：同系统审计保护级。

（2）强制访问控制：可信计算基对所有主体及其控制的客体（例如进程、文件、段和设备）实施强制访问控制。通过敏感标记为这些主体及客体指定安全等级。安全等级用二

维组表示：第一维是等级分类（如秘密、机密和绝密等），第二维是范畴（如适用范畴）。它们是实施强制访问控制的依据。可信计算基支持两种或两种以上成分组成的安全级。可信计算基控制的所有主体对客体的访问应满足以下要求。

① 仅当主体安全级中的等级分类高于或等于客体安全级中的等级分类，且主体安全级中的非等级类别包含了客体安全级中的全部非等级类别，主体才能读客体。

② 仅当主体安全级中的等级分类低于或等于客体安全级中的等级分类，且主体安全级中的非等级类别包含了客体安全级中的全部非等级类别，主体才能写一个客体。

可信计算基使用身份和鉴别数据来鉴别用户的身份，并保证用户创建的可信计算基外部主体的安全级和授权受该用户的安全级和授权的控制。

（3）敏感标记：是实施强制访问的基础。可信计算基应明确规定需要标记的客体（例如进程、文件、段和设备），明确定义标记的粒度（如文件级、字段级等），并必须使其主要数据结构具有相关的敏感标记。为了输入未加安全标记的数据，可信计算基向授权用户要求并接受这些数据的安全级别，且可由计算机信息系统可信计算基审计。

（4）身份鉴别：可信计算基初始执行时，首先要求用户标识自己的身份，而且，可信计算基维护用户身份识别数据并确定用户的访问权及授权数据。其他同系统审计保护级。

（5）客体重用：同系统审计保护级。

（6）审计：在系统审计保护级的基础上，要求可信计算基具有审计更改可读输出记号的能力。

（7）数据完整性：可信计算基通过自主和强制完整性策略，阻止非授权用户修改或破坏敏感信息。在网络环境中，使用完整性敏感标记来确保信息在传送中未受损。

**4. 第四级：结构化保护级**

本级的计算机信息系统可信计算基建立于一个明确定义的形式化安全策略模型之上，将第三级系统中的自主和强制访问控制扩展到所有主体与客体。此外，还要考虑隐蔽信道。本级的可信计算基必须结构化为关键保护元素和非关键保护元素；可信计算基的接口也必须明确定义，使其设计与实现能经受更充分的测试和更完整的复审；加强了鉴别机制；支持系统管理员和操作员的职能；提供可信设施管理；增强了配置管理控制。系统具有相当的抗渗透能力。

与安全标记保护级相比，本级的主要特征如下。

（1）可信计算基基于一个明确定义的形式化安全保护策略。

（2）将第三级实施的（自主或强制）访问控制扩展到所有主体和客体，即在自主访问控制方面，可信计算基应维护由外部主体能够直接或间接访问的所有资源（例如主体、存储客体和输入输出资源）实施强制访问控制，为这些主体及客体指定敏感标记，这些标记是等级分类和非等级类别的组合，它们是实施强制访问控制的依据。

（3）审计。

① 同系统安全标记保护级。

② 计算机信息系统可信计算基能够审计利用隐蔽存储信道时可能被使用的事件。

（4）数据完整性：计算机信息系统可信计算基通过自主和强制完整性策略，阻止非授权

用户修改或破坏敏感信息。在网络环境中，使用完整性敏感标记来确保信息在传送中未受损。

（5）隐蔽信道分析：系统开发者应彻底搜索隐蔽存储信道，并根据实际测量或工程估算确定每一个被标识信道的最大带宽。

（6）可信路径：对用户的初始登录和鉴别，计算机信息系统可信计算基在它与用户之间提供可信通信路径，该路径上的通信只能由该用户初始化。

**5. 第五级：访问验证保护级**

本级的可信计算基满足引用监视器需求。访问监控器仲裁主体对客体的全部访问。访问监控器本身是抗篡改的，必须足够小，能够分析和测试。

为了满足访问监控器的需求，可信计算基在其构造时，排除那些对实施安全策略来说并非必要的代码；在设计和现实时，从系统工程角度将其复杂性降低到最小程度。支持安全提供系统恢复机制管理员职能；扩充审计机制，当发生与安全相关的事件时发出信号；提供系统恢复机制。系统具有很高的抗渗透能力。

与第四级相比，本级的主要特征有如下几个方面。

（1）可信计算基的构造方面：本级具有访问监控器。访问监控器是监视主体和客体之间授权关系的部件，仲裁主体对客体的全部访问。访问监控器必须是抗篡改的，并且是可分析和测试的。

（2）在自主访问控制方面：由于有访问监控器，所以访问控制能为每个客体指定用户和用户组，并规定他们对客体的访问模式。没有存储权的用户只允许由授权用户指定对客体的访问权。

（3）在审计方面：可信计算基包含能够监控可审计安全事件发生与积累的机制，当超过阈值时，能够立即向安全管理员发出警报。并且，如果这些与安全相关的事件继续发生或积累，系统应以最小的代价终止它们。

（4）可信恢复：提供过程和机制，保证计算机信息系统失效或中断后，可以进行不损害任何安全保护性能的恢复。

### 5.5.3 信息安全测评认证体系

**1. 一般国家的信息安全测评认证体系**

目前世界许多国家都建立了国家信息安全测评认证体系。图 5.7 为已经建立 CC 信息安全测评认证体系的国家信息安全测评认证机构组织的一般结构。

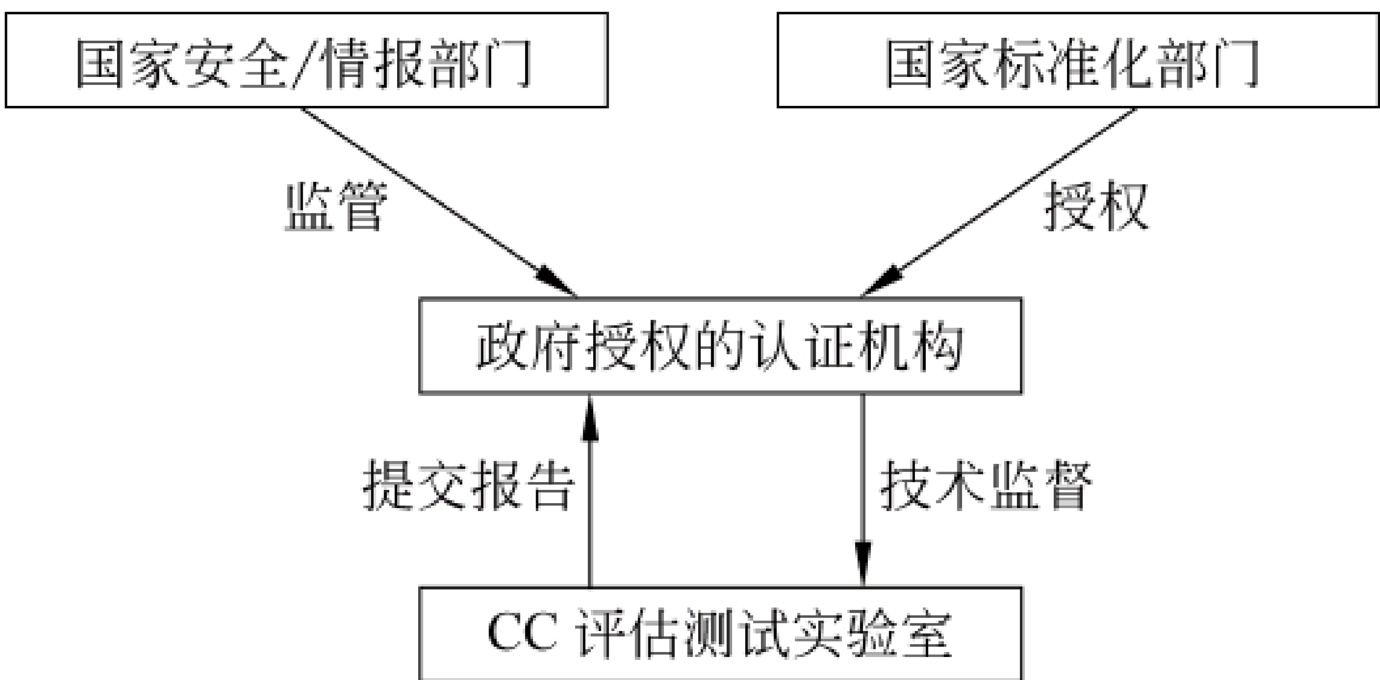



图 5.7　国家信息安全测评认证机构组织的一般结构

在这样的安全测评认证组织结构中，认证机构是核心。认证机构是一些公正的第三方，负责具体管理信息安全产品的安全性评估和认证，并颁发认证证书。它们上由国家标准化部门认可和授权，并受国家安全和情报主管部门的监管；下可委托一些具有商业性质的CC测试实验室进行安全性评估和认证的具体实施，并向认证机构提交结果。

**2. 国际互认**

1995年CC项目组成立了CC国际互认工作组，并于1997年制定了过渡性互认协定。目前，参加CC互认协定（CCRA）已经有美国的NSA和NIST、加拿大的CSE、英国的CESG、德国的GISA、法国的SCSSI、新西兰的DSD，以及澳大利亚、荷兰、西班牙、意大利、挪威、芬兰、瑞典、希腊等20多个国家的政府官方组织。目前CCRA也已经允许有政府机构参与或授权的非官方组织参加。

**3. 中国国家信息安全测评认证中心**

中国国家信息安全测评认证中心是国家授权的、并按照CC准则建立的具有第三方性质的技术机构。它代表国家，并依照国家认证的法律、法规和信息安全管理政策，对信息技术、信息系统、信息安全产品以及安全服务的安全性实施测试、评估和认证。图5.8为中国国家信息安全测评认证中心开展涉密信息系统认证的流程。

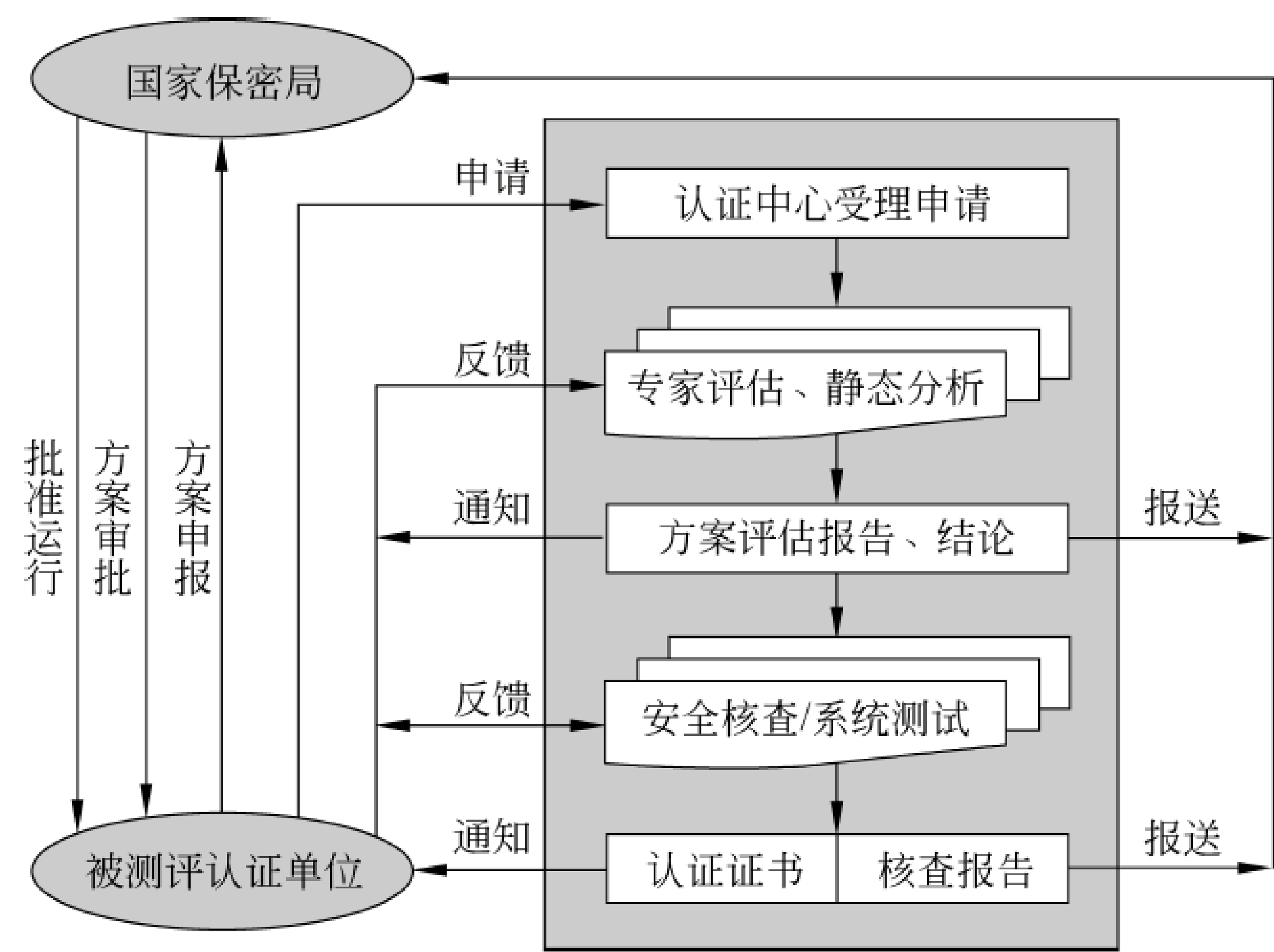



图5.8　中国国家信息安全测评认证中心开展涉密信息系统认证的流程

## 5.6　开放系统互连安全体系结构

一个信息系统的安全涉及有关安全的众多因素和要求，如保密性、完整性、可扩展性、方便性以及成本等。这些要求往往是有冲突的。特别重要的是，存在安全理论与系统实际之间的特殊性冲突。因此一个系统安全的最后实现，一定是这些众多因素的协调和折中考虑，也是理论如何应用于实际的问题。研究信息系统安全体系结构的目的，就是将普遍性

的安全体系原理与信息系统自身的实际相结合，从所需要保护的信息系统资源出发，分析由系统可能的威胁和系统自身可能的漏洞形成的安全风险，建立系统安全需求，从管理和技术上保证安全策略得以完整准确的实现，安全需求得以全面准确的满足。

本节介绍开放系统互连安全体系结构，即著名的 ISO/OSI 安全体系结构。它是国际标准化组织 ISO 于 1989 年在对 OSI（开放系统互连）环境的安全性进行深入研究的基础上提出的 ISO 7498-2 和《Internet 安全体系结构》（RFC 2401）。中国国家标准《信息处理系统开放系统互联基本参考模型——第二部分：安全体系结构》（GB/T 9387.2—1995）是一个与之等同的标准，它给出了基于 OSI 参考模型七层协议之上的信息安全体系结构。

### 5.6.1 开放系统互连安全体系结构概述

OSI 安全体系结构是一个普遍适用的安全体系结构，其核心内容是保证异构计算机系统进程与进程之间远距离交换信息的安全；其基本思想是，为了全面而准确地满足一个开放系统的安全需求，必须在 7 个层次中提供必需的安全服务、安全机制和技术管理，以及它们在系统上的合理部署和关系配置。这个体系结构可以用图 5.9 表示。

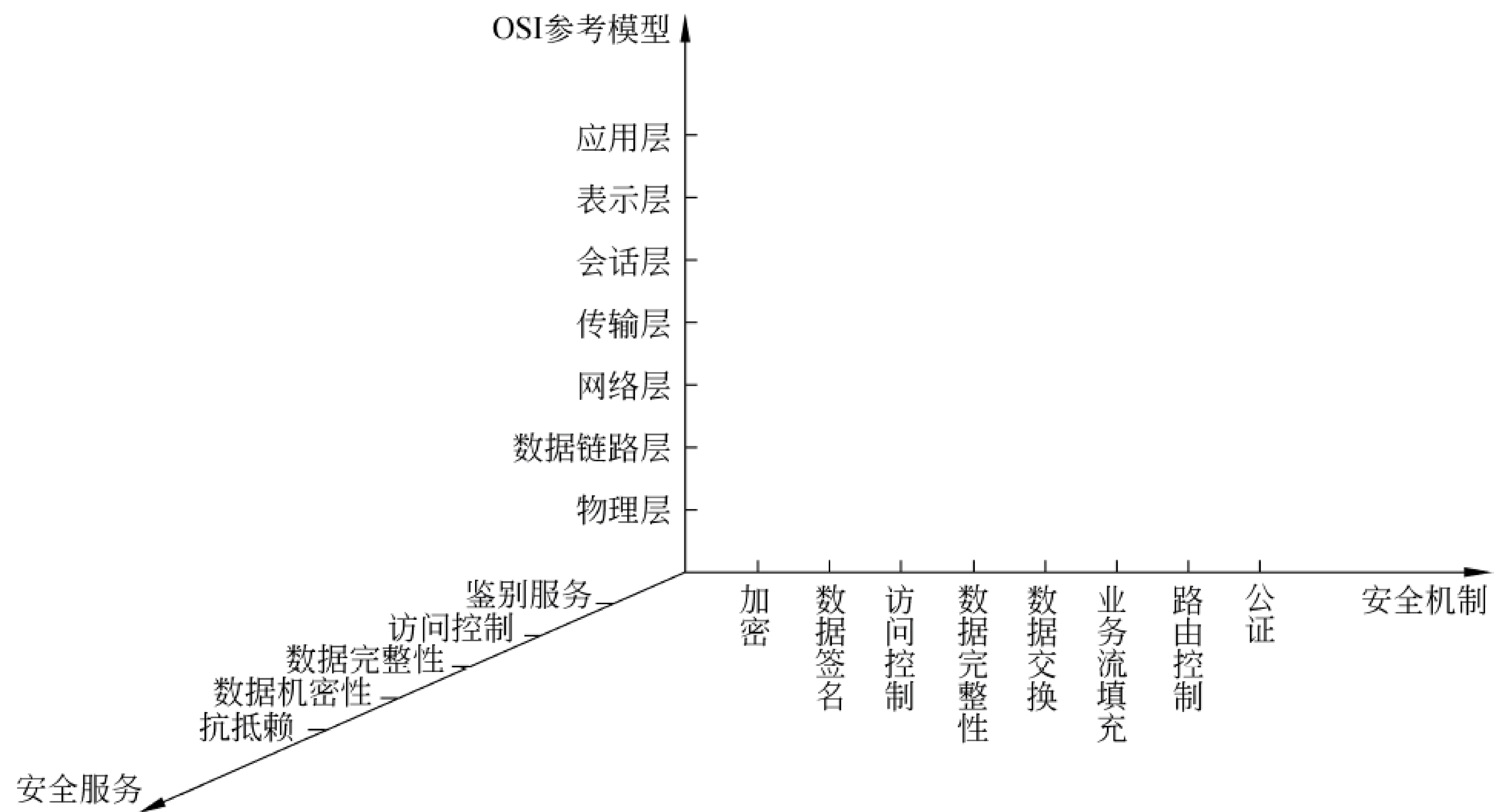



图 5.9 OSI 安全体系结构

OSI 安全体系结构提供的内容如下。

（1）提供安全体系结构所配备的安全服务（也称为安全功能）和有关安全机制在体系结构下的一般描述。

（2）确定体系结构内部可以提供相关安全服务的位置。

（3）保证完全准确地配置安全服务，并且一直维持于信息系统安全的生命周期中，安全服务必须满足一定强度的要求。

（4）一种安全服务可以通过某种单独的安全机制提供，也可以通过多种安全机制联合提供，一种安全机制可用于提供一种或多种安全服务，在七层协议中除第五层（会话层）外，每一层均能提供相应的安全服务。

实际上，最适合配置安全服务的是在物理层、网络层、传输层和应用层上。其他层都

不宜配置安全服务。

### 5.6.2 OSI 安全体系结构的安全服务

OSI 安全体系结构中定义的 5 大类安全服务，也称为安全防护措施。

#### 1. 鉴别服务

鉴别是最基本的安全服务，是对付假冒攻击的有效方法。鉴别可以分为对等实体鉴别和数据源鉴别。

1）对等实体鉴别

对等实体鉴别是在开放系统的两个同层对等实体间建立连接和传输数据期间，为证实一个或多个连接实体的身份而提供的一种安全服务。这种服务可以是单向的，也可以是双向的；可以带有有效期检验，也可以不带。从七层参考模型看，当由 *N* 层提供这种服务时，将使 *N*+1 层实体确信与之打交道的对等实体正是它所需要的对等 *N*+1 层实体。

2）数据源鉴别

数据源鉴别服务是对数据单元的来源提供识别，但对数据单元的重复或篡改不提供鉴别保护。从七层参考模型看，当由 *N* 层提供这种服务时，将使 *N*+1 层实体确信数据来源正是它所需要的对等 *N*+1 层实体。

#### 2. 访问控制

访问控制用于防止资源未授权使用。在 OSI 安全体系结构中，访问控制的安全目标如下。

（1）通过进程（可以代表人员或其他进程行为）对数据不同进程或其他计算资源的访问控制。

（2）在一个安全域内的访问或跨越一个或多个安全域的访问控制。

（3）按照其上下文进行的访问控制。如根据试图访问的时间、访问者地点或访问路由等因素的访问控制。

（4）在访问期间对授权更改做出反应的访问控制。

#### 3. 机密性

机密性就是保护信息（数据）不泄露或不泄露给那些未授权掌握这一信息的实体。

在信息系统安全中需要区分两类机密性服务。

（1）数据机密性服务：使攻击者想要从某个数据项中推出敏感信息是十分困难的。

（2）业务流机密性服务：使攻击者想要通过观察通信系统的业务流来获得敏感信息是十分困难的。

根据所加密的数据项，机密性服务可以有如下几种类型。

（1）连接机密性：为一次连接上的所有用户数据提供机密性保护。

（2）无连接机密性：为单个无连接的 SDU 中的全部用户数据提供机密性保护。

（3）选择字段机密性：为那些被选择的字段提供机密性保护。这些字段或处于连接的

用户中，或者为单个无连接的 SDU 中的字段。

（4）通信业务流机密性：使得通过观察通信业务流而不可能推断出其中的机密信息。

### 4. 完整性

完整性服务用于对抗数据在存储、传输等处理过程中受到的非授权修改，可分为 3 种重要类型。

（1）连接完整性服务。

（2）无连接完整性服务。

（3）选择字段完整性服务。

完整性服务还可以按是否具有恢复功能分为以下两种类型。

（1）不具有恢复功能的完整性服务。

（2）具有恢复功能的完整性服务。

OSI 安全体系把完整性服务概括为以下 5 个方面。

（1）带恢复的连接完整性：为连接上的所有用户数据保证其完整性，并检测整个 SDU 序列中的数据遭受到的任何篡改、插入和删除，或者同时进行补救和/或恢复。

（2）不带恢复的连接完整性：服务同带恢复的连接完整性，只是不做补救恢复。

（3）选择字段的连接完整性：为一次连接上传送的 SDU 的用户数据中的选择字段提供完整性保护，确定被选择字段是否遭受了篡改、插入、删除或不可用。

（4）无连接完整性：为单个无连接上的 SDU 提供完整性保护，检测一个接收到的 SDU 是否遭受了篡改，并在一定程度上提供对连接重放的检测。

（5）选择字段的无连接完整性：为单个无连接上的 SDU 中的选择字段提供完整性保护，检测被选择字段是否遭受了篡改。

### 5. 抗抵赖

上面的安全服务是针对来自未知攻击者的威胁，而抗抵赖服务的目的是保护通信实体免遭来自其他合法实体的威胁。OSI 定义的抗抵赖服务有两种类型。

（1）有数据原发证明的抗抵赖：为数据的接收者提供数据的原发证据，使发送者不能抵赖这些数据的发送或否认发送内容。

（2）有交付证明的抗抵赖：为数据的发送者提供数据交付证据，使接收者不能抵赖收到这些数据或否认接收内容。

## 5.6.3 OSI 七层中的安全服务配置

### 1. 安全分层及服务配置原则

安全服务分层以及安全机制在 OSI 七层上的配置应按照下列原则进行。

（1）实现一种服务的不同方法越少越好。

（2）在多层上提供安全服务来建立安全系统是可取的。

（3）为安全所需的附加功能不应该也不必要重复 OSI 的现有功能。

（4）避免破坏层的独立性。

（5）可信功能度的总量应尽量少。

（6）只要一个实体依赖于由位于较低层的实体提供的安全机制，那么任何中间层应该按不违反安全的方式构建。

（7）只要可能，就应以作为自容纳模块起作用的方法来定义一个层的附加安全功能。

（8）本标准被认定用于由包含所有 7 层的端系统组成的开放系统以及中继系统。

**2. 在 OSI 各层中的安全服务配置**

OSI 各层提供的安全服务配置如表 5.12 所示。不论所要求的安全服务是由该层提供还是由下层提供，各层上的服务定义都可能需要修改。

**表 5.12　OSI 各层中的安全服务配置**

| 安全服务 | 协议层 | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 对等实体鉴别 | | | √ | √ | | | √ |
| 数据源鉴别 | | | √ | √ | | | √ |
| 访问控制 | | | √ | √ | | | √ |
| 连接机密性 | √ | √ | √ | √ | | √ | √ |
| 无连接机密性 | | √ | √ | √ | | √ | √ |
| 连接字段机密性 | | | | | | | √ |
| 通信业务流机密性 | | | | | | √ | √ |
| 带恢复的连接完整性 | √ | | √ | | | | √ |
| 不带恢复的连接完整性 | | | | √ | | | √ |
| 选择字段连接完整性 | | | √ | √ | | | √ |
| 无连接完整性 | | | | | | | √ |
| 选择字段无连接完整性 | | | √ | √ | | | √ |
| 有数据原发证明的抗抵赖 | | | | | | | √ |
| 有交付证明的抗抵赖 | | | | | | | √ |

注：表中空白表示不提供。

### 5.6.4　OSI 安全体系结构的安全机制

OSI 安全体系结构没有说明 5 种安全服务如何实现，但是它给出了 8 种基本（特定的）安全机制，使用这 8 种安全机制，再加上几种普遍性的安全机制，将它们设置在适当的（*N*）层上，用以提供 OSI 安全体系结构安全服务。

**1．OSI 的 8 种特定的安全机制**

1）加密

在 OSI 安全体系结构的安全机制中，加密涉及 3 个方面的内容。

（1）密码体制的类型：对称密码体制和非对称密码体制。

（2）密钥管理。

（3）加密层的选取。表 5.13 给出了加密层选取时要考虑的因素，它不推荐在数据链路层上的加密。

**表 5.13　加密层选取时要考虑的因素**

| 加 密 要 求 | 加密层 |
| --- | --- |
| 对全部通信业务提供加密 | 物理层 |
| 细粒度保护（对每个应用提供不同的密钥）<br>抗抵赖或选择字段保护 | 表示层 |
| 提供机密性与不带恢复的完整性<br>对所有端对端之间通信的简单块进行保护<br>希望有一个外部的加密设备（如为了给算法和密钥提供物理保护或防止软件错误） | 网络层 |
| 提供带恢复的完整性以及细粒度保护 | 传输层 |

2）数字签名

数据签名是附加在数据单元上的一些数据，或是对数据单元所做的密码变换，这种附加数据或变换可以起如下作用。

（1）供接收者确认数据来源。

（2）供接收者确认数据完整性。

（3）保护数据，防止他人伪造。

数字签名需要确定两个过程。

（1）对数据单元签名，使用签名者私有（独有或机密的）信息。

（2）验证签过名的数据单元，使用的规程和信息是公开的，但不能推断出签名者的私有信息。

3）访问控制

访问控制是一种对资源访问或操作加以限制的策略。此外，它还可以支持数据的机密性、数据完整性、可用性以及合法使用的安全目标。访问控制机制可应用于通信联系中的任一端点或任一中间点。

访问控制机制可以建立在下面的一种或多种手段之上。

（1）访问控制信息库，保存了对等实体的访问权限。

（2）鉴别信息，如口令等。

（3）权限。

（4）安全标记。

（5）试图访问的时间。

（6）试图访问的路由。

（7）访问持续期。

4）数据完整性

数据完整性保护的目的是避免未授权的数据乱序、丢失、重放、插入和篡改。下面讨论数据完整性的两个方面：单个数据或字段的完整性和数据单元流或字段流的完整性。

决定单个数据单元的完整性涉及两个实体：一个在发送实体上，一个在接收实体上。发送实体给数据单元附上一个附加量，接收实体也产生一个相应的量，通过比较两者，可以判定数据在传输过程中是否被篡改。

对于连接方式数据传送，保护数据单元序列的完整性（包括防止乱序、数据丢失、重放或篡改），还需要明显的排序标记，如顺序号、时间标记或密码链；对于无连接数据传送，时间标记可以提供一定程度的保护，防止个别数据单元重放。

5）鉴别交换

（1）可用于鉴别交换的技术如下。

① 鉴别信息，如口令。

② 密码技术。

③ 使用该实体特征（生物信息等）或占有物（信物等）。

（2）可以结合使用的技术如下。

① 时间标记与同步时钟。

② 两次握手（单方鉴定）和三次握手（双方鉴定）。

③ 数字签名和公证。

6）通信业务填充

通信业务填充是一种反分析技术，通过虚假填充将协议数据单元达到一个固定长度。它只有受到机密服务保护才有效。

7）路由选择控制

路由选择控制机制可以使敏感数据只在具有适当保护级别的路由上传输，并且采取如下一些处理。

① 检测到持续的攻击，可以为端系统建立不同的路由的连接。

② 依据安全策略，使某些带有安全标记的数据禁止通过某些子网、中继或链路。

③ 允许连接的发起者（或无连接数据单元的发送者）指定路由选择，或回避某些子网、中继或链路。

8）公证

公证机制是由可信的第三方提供数据完整性、数据源、时间和目的地等的认证和保证。

**2．OSI 安全服务与安全机制之间的关系**

表 5.14 为 OSI 安全服务与安全机制之间的关系。

表 5.14 OSI 安全服务与安全机制之间的关系

| 安全服务 | 安全机制 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 加密 | 数字签名 | 访问控制 | 数据完整性 | 鉴别交换 | 业务填充 | 路由控制 | 公证 |
| 对等实体鉴别 | √ | √ | | | √ | | | |
| 数据源鉴别 | √ | √ | | | | | | |
| 访问控制 | | | √ | | | | | |
| 连接机密性 | √ | | | | | | √ | |
| 无连接机密性 | √ | | | | | | √ | |
| 连接字段机密性 | √ | | | | | | | |
| 流量机密性 | √ | | | | | √ | √ | |
| 带恢复的连接完整性 | √ | | | √ | | | | |
| 不带恢复的连接完整性 | √ | | | √ | | | | |
| 选择字段连接完整性 | √ | | | √ | | | | |
| 无连接完整性 | √ | √ | | √ | | | | |
| 选择字段无连接完整性 | √ | √ | | √ | | | | |
| 原发方抗抵赖 | | √ | | √ | | | | √ |
| 接收方抗抵赖 | | √ | | √ | | | | √ |

### 5.6.5 OSI 安全体系的安全管理

OSI 安全管理活动有如下 3 类：系统安全管理、安全服务管理和安全机制管理。此外还必须考虑 OSI 本身的安全和特定的系统安全管理活动。

#### 1. 系统安全管理

系统安全管理着眼于 OSI 总体环境的管理，其典型活动如下。

（1）总体安全策略的管理，包括一致性修改与维护。

（2）与别的 OSI 安全管理的相互作用。

（3）与安全服务管理和安全机制管理的交互。

（4）事件处理管理。在 OSI 中可以看到的是事件管理的实例，是远程报告的明显违反安全的企图以及对用来触发事件报告的阈值的修改。

（5）安全审计管理。主要内容有：

① 选择将被记录和被远程收集的事件。

② 授予或取消对所选事件进行审计跟踪日志记录的能力。

③ 所选审计记录的收集。

④ 准备安全审计报告。

（6）安全恢复管理。主要内容有：

① 维护用于对实有的或可疑的安全事件做出反应的规则。

② 远程报告明显的系统安全违规。

③ 安全管理者的交互。

### 2. 安全服务管理

安全服务管理指特定安全服务的管理。在管理一种特定安全服务时，可能的典型的活动如下。

（1）为该种服务决定并指派安全保护的目标。

（2）在有可选择的情况时，指定与维护选择规则。

（3）对需要事先取得管理者同意的安全机制进行协商。

（4）通过适当的安全机制管理功能，调用特定的安全机制。

（5）与其他安全服务管理功能和安全机制管理功能交互。

### 3. 安全机制管理

安全机制管理指特定安全机制的管理。典型的安全机制管理有下列一些。

1）密钥管理

（1）间歇性地产生与所要求的安全级别相称的合适密钥。

（2）根据访问控制的要求，决定每个密钥应分发给哪个实体。

（3）用可靠办法使这些密钥对开放系统中的实体是可用的，或将这些密钥分配给它们。

2）加密管理

（1）与密钥管理交互。

（2）建立密码参数。

（3）密码同步。

3）数字签名管理

（1）与密钥管理交互。

（2）建立密码参数与密码算法。

（3）在通信实体与可能有的第三方之间使用协议。

4）访问控制管理

（1）安全属性（包括口令）的分配。

（2）对访问控制表或权力表进行修改。

（3）在通信实体与其他提供访问控制服务的实体之间使用协议。

5）数据完整性管理

（1）与密钥管理交互。

（2）建立密码参数与密码算法。

（3）在通信实体间使用协议。

6）鉴别管理

（1）将说明信息、口令或密钥（使用密钥管理）分配给要求执行鉴别的实体。

（2）在通信的实体与其他提供鉴别服务的实体之间使用协议。

7）通信业务填充管理

（1）指定数据率。
（2）指定随机数据率。
（3）指定报文特性，例如长度等。
（4）可能按时间或日历来改变这些规定。

8）路由选择控制管理

路由选择管理的主要功能是确定那些按特定准则被认为是安全可靠和可信任的链路或子网络。

9）公证管理

（1）分配有关公证的信息。
（2）在公证方与通信的实体之间使用协议。
（3）与公证方的交互作用。

**4．OSI 管理的安全**

所有 OSI 管理功能的安全以及 OSI 管理信息的通信安全是 OSI 安全的重要部分。这一类安全管理将对上面所列的 OSI 安全服务与机制进行适当的选取，以确保 OSI 管理协议与信息获得足够的保护。例如，在管理信息库的管理实体之间的通信一般要求某种形式的保护。

**5.** 特定的系统安全管理活动

1）事件处理管理

（1）远程报告违反系统安全的明显企图。
（2）对用来触发事件报告的阈值进行修改。

2）安全审计管理

（1）选择将被记录和被远程收集的事件。
（2）授予或取消对所选事件进行审计跟踪日志记录的能力。
（3）所选审计记录的远程收集。
（4）准备安全审计报告。

3）安全恢复管理

（1）维护那些用来对实有的或可疑的安全事故做出反应的规则。
（2）远程报告对系统安全的明显违反。
（3）安全管理者的交互作用。

# 习　题

## 一、选择题

1. 系统备份与普通数据备份的不同在于，它不仅备份系统中的数据，还备份系统中安装的应用程序、数据库系统、用户设置和系统参数等信息，以便迅速________。

A. 恢复整个系统　　B. 恢复所有数据　　C. 恢复全部程序　　D. 恢复网络设置

2. 灾难恢复计划或者业务连续性计划关注的是信息资产的________属性。

A. 可用性　　B. 真实性　　C. 完整性　　D. 保密性

3. 数据备份常用的方式主要有完全备份、增量备份和________。

A. 逻辑备份　　B. 按需备份　　C. 差分备份　　D. 物理备份

4. 审计管理是指________。

A. 保证数据接收方收到的信息与发送方发送的信息完全一致

B. 防止因数据被截获而造成的泄密

C. 对用户和程序使用资源的情况进行记录和审查

D. 信息使用者都可以得到相应授权的全部服务

5. 关于安全审计目的描述错误的是________。

A. 识别和分析未经授权的动作或攻击　　B. 记录用户活动和系统管理

C. 将动作归结到为其负责的实体　　D. 实现对安全事件的应急响应

6. 安全审计跟踪是________。

A. 安全审计系统检测并追踪安全事件的过程

B. 安全审计系统收集易于安全审计的数据的过程

C. 人利用日志信息进行安全事件分析和追溯的过程

D. 对计算机系统中的某种行为的详尽跟踪和观察

7. “保护数据库，防止因未经授权的或不合法的使用造成的数据泄露、更改、破坏。”这是指数据的________保护。

A. 安全性　　B. 完整性　　C. 并发　　D. 恢复

8. 信息安全评测标准CC是________标准。

A. 美国　　B. 国际　　C. 中国　　D. 加拿大

9. 我国《信息系统安全等级保护基本要求》中，对不同级别的信息系统应具备的基本安全保护能力进行了要求，共划分为________级。

A. 4　　B. 5　　C. 6　　D. 7

10. 在CC中，称为访问控制保护级别的是________。

A. C1　　B. B1　　C. C2　　D. B2

## 二、问答题

1. 简述紧急响应的意义。

2. 试述紧急响应服务在实现目的方面受哪些因素制约。

3. 如何制定紧急响应预案？

4. 尽可能多地列举一些安全事件。

5. 简述应急事件处理的基本流程。

6. 灾难恢复涉及哪些内容？

7. 灾难恢复涉及哪些技术？

8. 简述数据容错和数据容灾之间的联系与区别。

9. 简述数据备份在数据容错和数据容灾中的作用。

10. 简述各种数据备份技术的特点。

11. 简述各种数据备份策略的用途。

12. 收集国内外有关应急响应、数据容错或数字取证的网站信息，简要说明各网站的特点。

13. 收集国内外有关应急响应、数据容错或数字取证的最新动态。

14. 论述数字证据的特征。

15. 上网搜索，提交一份有关数字取证工具的报告。

16. 如何保证数字证据的安全？

17. 审计与入侵检测技术有什么关系？

18. 简述安全审计的作用。

19. 简述日志的作用和记录内容。

20. 审计与入侵检测有什么关联？

21. 收集国内外有关安全审计的网站信息，简要说明各网站的特点。

22. 收集国内外有关安全审计的最新动态。

23. 风险评估对于信息系统安全有什么意义？

24. 为一个组织的信息系统进行安全风险评估。

25. NAI公司开发了一个用于安全风险评估的扫描器CyberCop Scanner，试安装并使用该工具。

26. 为学生成绩管理系统设计一个安全策略。这个系统最少要由学生、教师和管理人员访问。

27. 在一个具有读、写、准许和取消4种访问操作的系统中，准许操作可以授予其他主体读和写的访问权限，并且还可以授予其他主体发布对你拥有的资源的访问权限。如果你要使用准许和取消操作来控制对你所拥有的一个客体的所有访问，应当采用什么样的数据结构和算法来实现准许和取消操作？

28. 给出一个中等规模的局域网（包含一些子网，但不跨多个地域），为其设计一个安全解决方案。

29. 分析一个具体系统的安全需求，并给出相应的安全策略。

30. 如何理解OSI安全体系的安全机制和安全服务之间的对应关系？

31. 什么是可信计算基？

32. 详细说明安全标记保护级的可信计算基的功能。

33. 结构化保护级的主要特征有哪些？

34. 收集有关信息安全的定义、标准等方面的最新概念和进展。可以从下面的网站开始：

http://www.radium.ncsc.mil/tpep/process/fag.html

http://www.itsec.gov.uk

http://www.cse-cst.gc.ca/pub/criteria/CTCPE

35. 收集资料，分别给定出下列操作系统的安全等级，并说明理由。

（1）DOS。

（2）Windows。

（3）UNIX。

（4）Linux。

36. 写出以下与信息安全相关的英文缩写的全称：

FTP，HTTP，NFS，DNS，UDP，TCP，IP；

ICMP，SMTP，SNMP，POP，IMAP；

S/MIME，TFTP，TELNET；

Rlogin，RPC，NFS，NIS；

DCOM，ISN，SRC；

SYN，ACK，RST，FIN。

# 参 考 文 献


[1] 张基温. 信息系统安全教程[M]. 2 版. 北京：清华大学出版社，2015.

[2] 张基温. 信息系统安全原理[M]. 北京：中国水利水电出版社，2005.

[3] 张基温. 信息安全实验与实践教程[M]. 北京：清华大学出版社，2005.

[4] 张基温，江森林. 利用 Honeyd 诱测蠕虫的研究[J]. 江南大学学报（自然科学版），2005，(8).

[5] 张基温，蒋中云. 计算机取证概述[J]. 计算机教育，2005(10)：62-65.

[6] 张基温，陶利民. 一种基于移动 Agent 的新型分布式入侵检测系统[J]. 微计算机应用，2004，25(1).

[7] 陶利民，张基温. 轻量级网络入侵检测系统——Snort 的研究[J]. 计算机应用研究，2004(4)：106-108.

[8] 江森林，张基温. Honeyd 解析[J]. 计算机工程与设计，2005，26(3)：682-685.

[9] 王玉斐，张基温. 基于 NIDS 数据源的网络攻击事件分类技术研究[J]. 计算机应用，2005(12)：2748-2750.

[10] 蒋中云，张基温. 基于 Multi-Agent 的网络入侵取证模型的设计[J]. 微计算机信息，2005(12).

[11] 魏士靖，张基温. 基于犯罪画像的计算机取证分析方法研究[J]. 微计算机信息，2006(2).

[12] 张基温，王玉斐. 基于应用环境的入侵检测系统测试方案[J]. 计算机工程与设计，2006(7)：1220-1223.

[13] 叶茜，张基温. 基于移动代理的分布式拒绝服务攻击防御模型[J]. 计算机应用，2006(7)：1646-1648.

[14] 朱剑，张基温. 基于加权模糊推理的电子取证入侵重构系统[J]. 计算机工程与设计，2006(14)：2663-2665.

[15] 裴浩，张基温，黄可望. 基于 PMI 的 Web Service 访问控制方案[J]. 计算机工程与设计，2007(1)：59-61.

[16] 张基温，董瑜. 大规模 P2P 网络下蠕虫攻击的研究[J]. 微计算机信息，2008(3)：245-246.

[17] 张基温，刘英戈，陈广良，等. 基于 Mobile Agent 的协作式反垃圾邮件系统设计[J]. 计算机应用，2006，26(10)：2338-2340.

[18] 中国信息安全产品测评认证中心. 信息安全工程与管理[M]. 北京：人民邮电出版社，2003.

[19] 杨波. 现代密码学[M]. 北京：清华大学出版社，2003.

[20] 杜彦辉，等. 信息安全技术教程[M]. 北京：清华大学出版社，2013.

[21] 付永钢. 计算机信息安全技术[M]. 北京：清华大学出版社，2012.

[22] 彭新光，王峥. 信息安全技术与应用[M]. 北京：人民邮电出版社，2013.

[23] 张基温. 计算机网络原理[M]. 北京：高等教育出版社，2003.

[24] 张基温. 计算机网络技术[M]. 北京：高等教育出版社，2004.

[25] 黄波，刘洋洋. 信息安全法律汇编与案例分析[M]. 北京：清华大学出版社，2012.

[26] http://www.i023.com/.

[27] http://www.yesky.com/.

[28] http://www.anmeng.com.cn/page1/others/zjjt33.ph/.

[29] http://www.ca.jc.com/maindoc/.

[30] http://www.infosec.cs.pku.edu.cn/course/.

[31] http://bbs.yesky.com/book/html/1688.html/.

[32] http://www.hacker-defence.com/.

[33] https://alerts.securityfocus.com/.

[34] http://www.cisco.com/warp/public/.
[35] http://www.e8oo.com.cn/.
[36] http://www.sans.org/.
[37] http://www.networkingunlimited.com/white007.html/.
[38] http://us.cns911.com/holes/router/.
[39] http://www.securiteam.com/.

# 高等教育质量工程信息技术系列示范教材

## 系列主编：张基温

- 新概念 C 程序设计大学教程（第 4 版） 张基温
- 新概念 C++程序设计大学教程（第 3 版） 张基温
- 新概念 Java 程序设计大学教程（第 3 版） 张基温
- 计算机组成原理教程（第 8 版） 张基温
- 计算机组成原理解题参考（第 8 版） 张基温
- 计算机网络教程（第 2 版） 张基温
- 信息系统安全教程（第 3 版） 张基温
- 信息系统安全教程（第 3 版）习题详解 栾英姿
- 大学计算机——计算思维导论（第 2 版） 张基温
- UI 设计教程 牛金巍
- APP 开发教程——HTML5 应用 尹志军
- Python 大学教程 张基温